

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

<http://wrap.warwick.ac.uk/151121>

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk



Optimal Stopping and Randomisation

by

Matthew Zeng

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy in Statistics

Department of Statistics

January 2020

THE UNIVERSITY OF
WARWICK

Contents

Acknowledgments	iv
Declarations	v
Abstract	vi
Chapter 1 An Overview of the Thesis	1
Chapter 2 Classical Optimal Stopping Problem	6
2.1 All stopping times	6
2.2 Under finite stopping times	11
Chapter 3 Cautious Stochastic Choice, Optimal Stopping and Deliberate Randomisation	12
3.1 Introduction	12
3.2 Literature	13
3.3 CSC model in a static setting (Cerreia-Vioglio et al (2017))	14
3.4 The Optimal stopping models	16
3.4.1 Optimal stopping under expected utility	16
3.4.2 Optimal stopping under Cautious Stochastic Choice	20
3.5 A stylized example	21
3.5.1 Pure threshold strategies	23
3.5.2 Improvement with randomisation between Two Upper Thresholds	24
3.5.3 The Optimal Solution	26
3.5.4 A Generalized Example	31
3.6 Two realistic models	34
3.6.1 An example with S -shaped reference dependent utilities	35
3.6.2 An example based on concave utilities	39

3.7	Conclusion	42
Chapter 4 Optimal Stopping and the Sufficiency of Randomised Threshold Strategies		43
4.1	Introduction and main results	43
4.2	Problem specification and the problem in natural scale	45
4.3	Characterizing the possible laws of the stopped process in natural scale	46
4.4	Sufficiency of mixed threshold rules	48
4.4.1	An example - Rank dependent utility	49
4.5	Sufficient conditions for the optimality of pure threshold rules	50
4.6	Extension to other state spaces for the process in natural scale . . .	52
4.6.1	The range is unbounded below but bounded above	52
4.6.2	The range is bounded	52
4.6.3	The range is unbounded above and below	52
4.7	Conclusion	53
Chapter 5 Randomising Rules for Stopping Problems		55
5.1	Introduction	55
5.2	Problem Specification	56
5.3	The base case	58
5.4	Linear payoffs	65
5.4.1	Perceived continuation value $c = w$	65
5.4.2	Perceived continuation value $c = h$	66
5.4.3	Perceived continuation value $c = G$	66
5.4.4	Discussion	66
5.5	Call payoffs	68
5.5.1	Explicit formula for h	68
5.5.2	Perceived continuation value $c = w$	69
5.5.3	Perceived continuation value $c = h$	70
5.5.4	Perceived continuation value $c = G$	71
5.5.5	Comparison of the different solutions	71
5.5.6	Realised rate of stopping	74
5.6	Towards a model of continuous stopping	75
5.6.1	Modification of the randomising stopping rule	75
5.6.2	Making ξ dependent on λ	76
5.6.3	Alternative formulation of the limiting case	80
5.7	Conclusion	84

Chapter 6	Constrained Optimal Stopping, Liquidity and Effort	85
6.1	Introduction	85
6.2	The set-up	87
6.2.1	Some results for classical problems	89
6.3	Cost functions	91
6.4	Quadratic cost functions	91
6.4.1	Heuristics	91
6.4.2	First Example	92
6.4.3	Verification	104
6.5	Concave cost functions	109
6.5.1	An example	112
6.6	Numerical analysis on further examples	113
6.6.1	Addition of a linear cost	113
6.6.2	Quadratic costs with positive fixed cost	114
6.6.3	Cost functions defined on a subset of \mathbb{R}_+	117
6.7	Conclusion	121

Acknowledgments

I must first express my utmost gratitude to my supervisors David Hobson and Vicky Henderson for their excellent guidance during my PhD study. Apart from obtaining excellent academic advice, I also received invaluable support and encouragement over the past few years, which have been the crucial input into my professional development.

I also wish to acknowledge the financial support from a Chancellor's International Scholarship offered by Warwick, which made my study here possible.

Care from friends is unmatched. I would like to mention some who have played an important part in three years: Kaiye, Haodong, Shuo, Dominic, Jonathan, Takumi, Isabel, Chris.

My special thanks go to my parents for their unconditional love.

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. Work based on collaborative research is declared as follows:

Chapter 3 is a joint work with Vicky Henderson and David Hobson based on the paper "Cautious Stochastic Choice, Optimal Stopping, and Deliberated Randomisation", available on SSRN 3118906

Chapter 4 is a joint work with Vicky Henderson and David Hobson based on the paper "Optimal Stopping and the Sufficiency of Randomised Threshold Strategies", available on arXiv:1708.01038

Chapter 5 is a joint work with David Hobson based on the paper "Randomising Rules for Stopping Problems"

Chapter 6 is a joint work with David Hobson based on the paper "Constrained Optimal Stopping, Liquidity, and Effort", available on arXiv:1901.07270

Work conducted exclusively by myself is declared as follows:

- Chapter 2
- Chapter 3.3, Chapter 3.5.1, Chapter 3.5.2, Chapter 3.6
- Chapter 5.4, Chapter 5.5, Chapter 5.6.1, Chapter 5.6.2
- Chapter 6.3, Chapter 6.5, Chapter 6.6

Abstract

This thesis is a collection of four individual works on optimal stopping problems in junction with stochastic behaviours. Chapter 2 introduces the classical optimal stopping problem. In the classical model, the optimal strategy is to stop at some predetermined threshold, and thus there is no stochastic behaviours involved. Chapter 3 established a dynamic Cautious Stochastic Choice (CSC) model for an optimal stopping problem. Randomised strategies outperform threshold strategies in the CSC model, and thus, stochastic behaviours are predicted by our CSC model. Chapter 4 discussed the sufficiency of randomised threshold strategies and pointed out that the desire of stochastic behaviours stems from quasi-convexity. Chapter 5 considered a stopping problem where the agent doesn't stop with probability one. Instead, the stopping probability depends on the relative values of stopping and continuing. We discussed the case where stopping opportunities are constrained to be event times of an independent Poisson process. Dupuis and Wang introduced constraint on the class of admissible stopping times which they had to take values in the set of event times of an exogenous, time-homogeneous Poisson process. Chapter 6 extended the analysis of Dupuis and Wang (2005) to allow the agent to choose the rate of the Poisson process. Even for a simple model for the stopped process and a simple call-style payoff, the problem leads to a rich range of optimal behaviours which depend on the form of the cost function.

Chapter 1

An Overview of the Thesis

Stopping problems are often used to model dynamic decision-making tasks, such as option pricing, irreversible investment, market entry and job search, and are widely applied in finance, economics and statistics. Consider the classical optimal stopping problem. Let the asset price process $X = (X_t)_{t \geq 0}$ be a time-homogeneous, continuous, real-valued, strong-Markov process with initial value $X_0 = x$, living on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = \{\mathcal{F}_t\})$ which satisfies the usual conditions. Let $g : \mathbb{R} \mapsto \mathbb{R}_+$ be a (measurable) payoff function (satisfying suitable growth conditions, so that the problem is well-posed) and let β be a strictly positive discount factor. The value function $w = w(x)$ of the classical discounted optimal stopping problem is defined as

$$w(x) = \sup_{\tau \in \mathcal{T}} \mathbb{E}^x[e^{-\beta\tau} g(X_\tau)] \quad (1.1)$$

where \mathcal{T} is the set of all \mathbb{F} -stopping times and \mathbb{E}^x denotes the expectation conditioning on the initial state $X_0 = x$. In this paradigm, at each instant the agent makes a choice between stopping (and receiving an instantaneous payoff) and continuing (and receiving a discounted payoff in the future). Under optimising behaviour, the agent will stop if the stopping value is at least as large as the continuation value. One way to characterise w is via the variational inequality $\min\{\beta w - \mathcal{L}w, w - g\} = 0$ where $\mathcal{L} = \mathcal{L}^X$ is the infinitesimal generator of X . Typically, the optimal strategies are of threshold form (i.e. the stopping time is the first time the underlying process reaches some set).

Now, we consider the canonical case where X is an exponential Brownian motion started at x :

$$\frac{dX_t}{X_t} = \mu dt + \sigma dW_t; \quad X_0 = x \quad (1.2)$$

where μ and σ are positive constants. Then X has generator $\mathcal{L} = \mathcal{L}^X$ given by

$\mathcal{L}f = \frac{1}{2}x^2\sigma^2f'' + \mu xf'$. And let g be an American call option given by

$$g(x) = (x - K)_+. \quad (1.3)$$

The optimal stopping time in this case is of threshold type (i.e. stop as soon as the underlying process is above some predetermined level). More precisely, the optimal stopping time is given by (see Chapter 2 for more details)

$$\tau_{L^*} = \inf\{t > 0; X_t \geq L^*\}$$

where $L^* = \frac{\theta}{\theta-1}K$ and $\theta = (\frac{1}{2} - \frac{\mu}{\sigma^2}) + \sqrt{(\frac{1}{2} - \frac{\mu}{\sigma^2})^2 + \frac{2\beta}{\sigma^2}}$.

The stopping time τ_{L^*} is called a *threshold strategy* and it implies that if the agent is asked to choose an exercising price level for the same option multiple times, he always stops at the optimal threshold L^* given $x \leq L^*$ (for $x \geq L^*$, stopping is immediate). Unfortunately, this form of predicted behaviours from this classical optimal stopping problem does not match the observed behaviours in the financial markets where investors are seen to sell identical assets at different price levels. Furthermore, the phenomenon of stochastic choice has been consistently recorded by experimental studies. When subjects are asked to choose from the same set of options many times, they are inconsistent in their choices. Patterns of stochastic choice were first recorded by Tversky (1969) and many studies have replicated, explored and extended his results (see Agranov and Ortoleva (2017) for recent findings and a comprehensive overview, and amongst others Dwenger et al (2013), Hey and Orme (1994)). Hence, this thesis goes beyond the classical optimal stopping paradigm and builds dynamic, continuous-time models in which stochastic behaviours can be captured.

There is a large body of theoretical models which were developed to capture the phenomenon of stochastic or random choice. *Cautious Stochastic Choice* (CSC) falls into the class of stochastic models postulating that stochasticity is a *deliberate choice* of the agent.¹ Cerreia-Vioglio et al (2017) develop a theory of Cautious Stochastic Choice, showing that CSC agents may benefit from randomising in a static decision making setting. There are fewer models capturing the phenomenon of stochastic choice in the dynamic setting of a stopping problem. Chapter 3 of this thesis considers agents who face optimal timing decisions in a dynamic setting and who exhibit cautious stochastic choice. The CSC agent is unsure which utility

¹There are two other main classes of models of stochastic choice. In *random utility* models, subjects maximise a well defined utility function but this changes stochastically over time (eg. Gul and Pesendorfer (2006)). In models of *bounded rationality*, agents have well defined and stable preferences but may not make the best choice because of bounded rationality (see Johnson and Ratcliff (2013) for a review).

function to use from a family of possibilities \mathcal{W} and applies caution to choose the worst-case certainty equivalent. For a fixed stopping time τ and a fixed utility $u \in \mathcal{W}$, we define the corresponding certainty equivalent as $C_\tau^u = u^{-1}(\mathbb{E}[u(X_\tau)])$. The CSC value for a single stopping time τ is $V_\tau = \inf_{u \in \mathcal{W}} C_\tau^u$. Under CSC the optimal stopping problem is to find $V(\mathcal{S}) = \sup_{\tau \in \mathcal{S}} V_\tau$, that is

$$V(\mathcal{S}) = \sup_{\tau \in \mathcal{S}} \inf_{u \in \mathcal{W}} C_\tau^u \quad (1.4)$$

where \mathcal{S} is a set of stopping times. The optimal stopping rule is the one which maximises the CSC value. Our main result is that CSC agents may have an optimal strategy which is not of threshold form and may involve randomisation. Randomisation in this case is deliberate, being an optimal result derived from the dynamic CSC setting.

In light of the insufficiency of threshold strategies, in Chapter 4 we conduct investigations in the sufficiency of *randomised threshold strategies* (i.e. stopping rules which are based on the first exit from a randomly chosen interval). Our main result is that: for problems in which the value associated to a stopping rule depends on the law of the stopped process, if this value is quasi-convex on the space of attainable laws then it is sufficient to restrict attention to the class of threshold strategies which involves no randomisation; however, if the objective function is not quasi-convex, this may not be the case. We show that, nonetheless, it is sufficient to restrict attention to mixtures of threshold strategies.

The fact that quasi-convexity means that there is no benefit from following randomised strategies is well understood in the economics literature, see Machina (1985) Camerer and Ho (1994), Wakker (2010) and He et al (2017). Recently there has been a surge of interest in problems which, whilst they have the law invariance property, do not satisfy the quasi-convex criterion. Our dynamic CSC model is one example whose value function does not satisfy quasi-convex criterion. Also, the CSC value of a stopping rule depends only on the law of X_τ . By our main result, in searching for an optimal stopping rule for (1.4), it is sufficient to restrict attention to randomised threshold rules. By considering both stylised and realistic models in dynamic CSC setting, we can show randomised threshold strategies outperform cut-off or threshold strategies.

Apart from the non-quasi-convexity of the value function, stochastic behaviour can also stem from the stopping probability. In Chapter 5, we build a dynamic, continuous-time model of stopping in which the probability of stopping is not zero-one. Instead the probability of stopping depends on the relative values of the imme-

mediate receipts g and the perceived continuation value c (i.e. the stopping probability at any decision point is given by $p = \frac{g}{g+c}$). We call this a *randomised stopping rule*. In contrast to Chapter 3, randomisation is engineered into the setup of randomising stopping rule instead of being an optimal result.

One immediate issue of this *randomised stopping rule* is that if at each instant an agent has a positive probability of stopping, then since in a continuous-time model there are an uncountable number of stopping opportunities in any small interval, the agent will end up stopping immediately. To deal with this issue, we constrain the agent to stopping at a countable number of times, namely the event times $\{T_n^\lambda\}_{n \geq 1}$ of an independent Poisson process of rate λ . Optimal stopping problems in which stopping is only possible at event times of a Poisson process have been studied previously by Dupuis and Wang (2005). The value function for an optimal stopping problem with such constraints is given by

$$h(x) = h^\lambda(x) = \sup_{\tau \in \mathcal{T}^\lambda} \mathbb{E}^x[e^{-\beta\tau} g(X_\tau)] \quad (1.5)$$

where \mathcal{T}^λ is the set of all stopping times taking values in the event times of the Poisson process. Let $\hat{\mathcal{T}}_0^\lambda$ be the set of stopping times taking values in $\{0\} \cup \{T_n^\lambda\}_{n \geq 1}$. Let $V^{\lambda, h^\lambda} = V^h$ be the value of the optimal stopping problem, conditional on there being an event of the Poisson process at time 0. Then we have

$$V^h(x) = \sup_{\tau \in \hat{\mathcal{T}}_0^\lambda} \mathbb{E}^x[e^{-\beta\tau} g(X_\tau)] = \max\{g(x), h(x)\}. \quad (1.6)$$

Further, by conditioning on the first event time of the Poisson process we have

$$h(x) = \mathbb{E}^x \left[\int_0^\infty dt \lambda e^{-\lambda t} e^{-\beta t} V^h(X_t) \right]. \quad (1.7)$$

Substituting (1.6) and (1.7) gives an expression for h in feedback form.

$$h(x) = \mathbb{E}^x \left[\int_0^\infty dt \lambda e^{-\lambda t} e^{-\beta t} \{g(X_t) \vee h(X_t)\} \right]. \quad (1.8)$$

Chapter 5 considers the stopping problem where stopping can only occur at event times of a Poisson process, but in contrast to an optimal stopping problem where stopping probability is either zero or one, the probability of stopping under the randomising stopping rule depends on the value of immediate stopping $g = g(X_t)$ and on the perceived continuation value $c = c(X_t)$. More precisely, we suppose there

is a map $\Gamma : \mathbb{R}_+ \times \mathbb{R}_+ \mapsto [0, 1]$ such that the probability of stopping at time t is

$$p(X_t) = \Gamma(g(X_t), c(X_t)).$$

Integrating against the time of the first event of the Poisson process, and by analogy with (1.7), we obtain the continuation value function

$$G(x) = \mathbb{E}^x \left[\int_0^\infty dt \lambda e^{-\lambda t} e^{-\beta t} V^\Gamma(X_t) \right] \quad (1.9)$$

where V^Γ is the value function conditional on there being an event of the Poisson process at time zero. Using the randomising stopping rule Γ , we get

$$V^\Gamma = \Gamma(g, c)g + (1 - \Gamma(g, c))G.$$

In Chapter 6, we extend the analysis of Dupuis and Wang (2005) to allow the agent to choose the rate of the Poisson process $\Lambda = (\Lambda_t)_{t \geq 0}$. Also, some cost is incurred per unit time : $C_t = C(\Lambda_t)$.

In this paradigm, the agent has two problems to solve. Firstly, for a given a rate of the Poisson process he needs to find the optimal stopping strategy. In analogy to (1.5), we write the value function for a given control Λ

$$H^\Lambda(x) = \sup_{\tau \in \mathcal{T}^\Lambda} \mathbb{E}^x \left[e^{-\beta \tau} g(X_\tau) - \int_0^\tau e^{-\beta s} C(\Lambda_s) ds \right]. \quad (1.10)$$

Secondly, the agents needs to work out the optimal control on the rate of the Poisson process so that the value is maximised. Our objective function is then defined as

$$H(x) = \sup_{\Lambda \in \mathcal{A}} H^\Lambda(x) \quad (1.11)$$

where \mathcal{A} is the admissible control space.

We interpret the rate of the Poisson process $\Lambda = (\Lambda_t)_{t \geq 0}$ as an indicator of how much effort the agent spends on searching for stopping opportunities. The harder the agent works, the larger the rate of the Poisson process gets. Our focus is on the case where X is an exponential Brownian motion, but the general case of a regular, time-homogeneous diffusion can be reduced to this case at the expense of slightly more complicated technical conditions. We begin by rigorously stating the form of the problem we will study. Then we proceed to solve for the effort process and stopping rule in (6.1). It turns out that there are two distinctive cases depending on the shape of C or more precisely on the finiteness of $\lim_{\lambda \uparrow \infty} \frac{C(\lambda)}{\lambda}$.

Chapter 2

Classical Optimal Stopping Problem

2.1 All stopping times

In this section we keep the setups defined by (1.1), (1.2), and (1.3). Moreover, we require $\lim_{t \rightarrow \infty} \mathbb{E}^x[e^{-\beta t} g(X_t)] = 0$ so that the problem is well defined for infinite stopping times. Using that fact that $g(x) \leq x$ and the properties of geometric Brownian motion we have

$$0 \leq \lim_{t \rightarrow \infty} \mathbb{E}^x[e^{-\beta t} g(X_t)] \leq \lim_{t \rightarrow \infty} \mathbb{E}^x[e^{-\beta t} X_t] \leq \lim_{t \rightarrow \infty} x e^{(\mu - \beta)t}$$

Thus, if we assume that $\mu < \beta$, it follows that $\lim_{t \rightarrow \infty} \mathbb{E}^x[e^{-\beta t} g(X_t)] = 0$. We characterise w via the variational inequality $\min\{\beta w - \mathcal{L}w, w - g\} = 0$ where $\mathcal{L}f = \frac{1}{2}x^2\sigma^2 f'' + \mu x f'$. The optimal stopping time in this case is conjectured to be a threshold type stopping rule (i.e. stop as soon as the underlying process is above some predetermined level). By (1.3), we know that it's better to continue when X is small. Therefore, we assume (and prove later) that the optimal stopping time is

$$\tau_L = \inf\{t > 0; X_t \geq L\} \tag{2.1}$$

Under this stopping rule, the state space is divided into a stopping region $\mathcal{S} = [L, \infty)$ and a continuation region $\mathcal{C} = (0, L)$. Denote by v_L the value of the option under strategy τ_L . On stopping region \mathcal{S} , v_L satisfies

$$v_L = g \tag{2.2}$$

On continuation region \mathcal{C} , v_L satisfies

$$\mathcal{L}v_L - \beta v_L = 0 \quad (2.3)$$

The general solution to (2.3) is given by $v_L(x) = Ax^\theta + A_0x^{\theta_0}$, where

$$\theta = \left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right) + \sqrt{\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + \frac{2\beta}{\sigma^2}} > 1 \quad (2.4)$$

$$\theta_0 = \left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right) - \sqrt{\left(\frac{1}{2} - \frac{\mu}{\sigma^2}\right)^2 + \frac{2\beta}{\sigma^2}} < 0. \quad (2.5)$$

Since we require $v_L(0) = 0$, A_0 must be 0. Moreover, v_L is continuous at $x = L$. Thus, we have

$$AL^\theta = L - K$$

which gives $A = \frac{L-K}{L^\theta}$. Hence, we expect

$$v_L(x) = \begin{cases} \left(\frac{x}{L}\right)^\theta (L - K), & x < L \\ (x - K)_+, & x \geq L \end{cases}.$$

The optimal threshold L^* maximises v_L . We set $\frac{\partial}{\partial L} \left\{ \left(\frac{x}{L}\right)^\theta (L - K) \right\} = 0$. Therefore, the optimal threshold L^* satisfies

$$\frac{L^* - K}{(L^*)^\theta} \theta (L^*)^{\theta-1} = 1 \quad (2.6)$$

and hence

$$\begin{cases} L^* = \frac{\theta}{\theta-1} K \\ A^* = \frac{(L^*)^{1-\theta}}{\theta} \end{cases} \quad (2.7)$$

It follows that under the optimal threshold L^* , the value function v for the call option is given by

$$v(x) = v_{L^*}(x) = \begin{cases} \frac{(\frac{\theta}{\theta-1}K)^{1-\theta}}{\theta} x^\theta, & x \in (0, \frac{\theta}{\theta-1}K) \\ (x - K)_+, & x \in [\frac{\theta}{\theta-1}K, \infty) \end{cases} \quad (2.8)$$

Note that $0 \leq v'(x) \leq 1$ and $v(x) \leq x$.

Theorem 1. *Suppose the underlying process is a geometric Brownian motion defined in (1.2) and let v be given by (2.8). Then $v = w$ where w is given by (1.1) and hence is the price of a perpetual American call option g defined in (1.3) ; the*

stopping time τ_L defined in (2.1) for $L = L^*$ is optimal.

Proof. (a) First we prove that v provides an upper bound for the value of this call option (1.1). For fixed $T > 0$ and for every stopping time $\tau \in \mathcal{T}$, applying Itô's formula to $e^{-\beta t}v(X_t)$ between 0 and $\tau \wedge T$ we have

$$e^{-\beta(\tau \wedge T)}v(X_{\tau \wedge T}) = v(x) + \int_0^{\tau \wedge T} e^{-\beta t}(\mathcal{L}v - \beta v)(X_t)dt + \int_0^{\tau \wedge T} e^{-\beta t}\sigma X_t v'(X_t)dW_t.$$

Using the the fact that $0 \leq v' \leq 1$ and the property of geometric Brownian motion we have

$$\begin{aligned} \mathbb{E}^x \left[\int_0^T (e^{-\beta t}\sigma X_t v'(X_t))^2 dt \right] &\leq \mathbb{E}^x \left[\int_0^T (e^{-\beta t}\sigma X_t)^2 dt \right] \\ &= \frac{\sigma^2 x^2 \left[1 - e^{(2\mu - 2\beta + \sigma^2)T} \right]}{2\beta - 2\mu - \sigma^2} < \infty. \end{aligned} \quad (2.9)$$

Thus, the stopped stochastic integral is a true martingale. Taking expectation on both sides we get

$$\mathbb{E}^x \left[e^{-\beta(\tau \wedge T)}v(X_{\tau \wedge T}) \right] = v(x) + \mathbb{E}^x \left[\int_0^{\tau \wedge T} e^{-\beta t}(\mathcal{L}v - \beta v)(X_t)dt \right].$$

Since the transition density of X is absolutely continuous with respect to Lebesgue measure and it can be checked that $\mathcal{L}v - \beta v \leq 0$ holds for every $x \in (0, \infty)$, it follows that

$$v(x) \geq \mathbb{E}^x \left[e^{-\beta(\tau \wedge T)}v(X_{\tau \wedge T}) \right].$$

Since $v \geq 0$, we apply Fatou's lemma and obtain,

$$\begin{aligned} v(x) &\geq \liminf_{T \rightarrow \infty} \mathbb{E}^x \left[e^{-\beta(\tau \wedge T)}v(X_{\tau \wedge T}) \right] \\ &\geq \mathbb{E}^x \left[\liminf_{T \rightarrow \infty} e^{-\beta(\tau \wedge T)}v(X_{\tau \wedge T}) \right] \\ &= \mathbb{E}^x \left[e^{-\beta\tau}v(X_\tau) \right] \geq \mathbb{E}^x \left[e^{-\beta\tau}g(X_\tau) \right] \end{aligned}$$

Since the above inequality holds for every $\tau \in \mathcal{T}$, we derive

$$v(x) \geq \sup_{\tau \in \mathcal{T}} \mathbb{E}^x \left[e^{-\beta\tau}g(X_\tau) \right] = w(x). \quad (2.10)$$

(b) Next, we want to show that v is a lower bound for this call option value (1.1).

For fixed $T > 0$, applying Itô's formula to $e^{-\beta t}v(X_t)$ between 0 and $\tau_{L^*} \wedge T$ we have

$$\begin{aligned} e^{-\beta(\tau_{L^*} \wedge T)}v(X_{\tau_{L^*} \wedge T}) &= v(x) + \int_0^{\tau_{L^*} \wedge T} e^{-\beta t}(\mathcal{L}v - \beta v)(X_t)dt \\ &\quad + \int_0^{\tau_{L^*} \wedge T} e^{-\beta t}\sigma X_t v'(X_t)dW_t. \end{aligned} \quad (2.11)$$

Similar to (2.9) we can show that the stopped stochastic integral part is a true martingale. Moreover, we have that $\mathcal{L}v - \beta v = 0$ for any $x \in (0, L^*)$. It follows that

$$\int_0^{\tau_{L^*} \wedge T} e^{-\beta t}(\mathcal{L}v - \beta v)(X_t)dt = 0.$$

Taking expectations on both sides of (2.11) we obtain

$$\begin{aligned} v(x) &= \mathbb{E}^x \left[e^{-\beta(\tau_{L^*} \wedge T)}v(X_{\tau_{L^*} \wedge T}) \right] \\ &= \mathbb{E}^x \left[e^{-\beta\tau_{L^*}}g(X_{\tau_{L^*}})\mathbb{I}_{\{\tau_{L^*} < T\}} \right] + \mathbb{E}^x \left[e^{-\beta T}v(X_T)\mathbb{I}_{\{\tau_{L^*} \geq T\}} \right]. \end{aligned}$$

Note that

$$e^{-\beta T}v(X_T)\mathbb{I}_{\{\tau_{L^*} \geq T\}} \leq \left(\sup_{x \in [0, L^*]} v(x) \right) e^{-\beta T}\mathbb{I}_{\{\tau_{L^*} \geq T\}}$$

Therefore, we have

$$v(x) \leq \mathbb{E}^x \left[e^{-\beta\tau_{L^*}}g(X_{\tau_{L^*}})\mathbb{I}_{\{\tau_{L^*} < T\}} \right] + M e^{-\beta T}\mathbb{P}_{\{\tau_{L^*} \geq T\}}.$$

where $M = \sup_{x \in [0, L^*]} v(x)$ is constant. Let $T \rightarrow \infty$, and use monotone convergence theorem to obtain,

$$v(x) \leq \mathbb{E}^x \left[e^{-\beta\tau_{L^*}}g(X_{\tau_{L^*}})\mathbb{I}_{\{\tau_{L^*} < \infty\}} \right].$$

We also notice that

$$e^{-\beta\tau_{L^*}}g(X_{\tau_{L^*}})\mathbb{I}_{\{\tau_{L^*} = \infty\}} = \limsup_{t \rightarrow \infty} e^{-\beta t}g(X_t) = 0, \quad \mathbb{P} - a.s.$$

Hence, we have

$$v(x) \leq \mathbb{E}^x \left[e^{-\beta\tau_{L^*}}g(X_{\tau_{L^*}}) \right] \leq \sup_{\tau \in \mathcal{T}} \mathbb{E}^x \left[e^{-\beta\tau}g(X_\tau) \right] = w(x). \quad (2.12)$$

Combining (2.10) and (2.12), it completes the proof.

Remark 1. It is evident from the definition of τ_L and the fact X is a geometric

Brownian motion that τ_L is not always finite. Using the well-known Doob formula (see e.g. [45], Chap. VII)

$$\mathbb{P} \left[\sup_{t \geq 0} \left(W_t - \frac{1}{2}at \right) \geq b \right] = e^{-ab}$$

for a and b positive. It is straightforward to verify that

$$\mathbb{P}_x(\tau_L < \infty) = \mathbb{P}_x(\sup_{t \geq 0} X_t \geq L) = \begin{cases} 1 & \text{if } 2\mu \geq \sigma^2 \text{ or } x \in [L, \infty) \\ \left(\frac{x}{L}\right)^{1-2\mu/\sigma^2} & \text{if } 2\mu \leq \sigma^2 \text{ and } x \in (0, L) \end{cases}.$$

Remark 2. First, let's consider the case where $\beta < \mu$. The requirement $\lim_{t \rightarrow \infty} \mathbb{E}^x[e^{-\beta t}g(X_t)] = 0$ is not satisfied in this case. It is easy to see that $w(x) = \infty$ for all $x > 0$. Indeed, in this case, from the explicit expression of the geometric Brownian motion, we have for any $T > 0$,

$$w(x) \geq \mathbb{E}^x \left[e^{-\beta T} (X_T - K) \right] = xe^{(\mu-\beta)T} - Ke^{-\beta T}$$

and by sending T to infinity, we get the announced result. In this case, the underlying price process grows faster enough to compensate the loss from discounting. As a result, it's never optimal to stop.

For the case $\beta = \mu$, we have $w(x) = x$. Indeed, for any stopping time τ , for any $n \in \mathbb{N}$, by the optional sampling theorem for $e^{-\beta t}X_t$ we have

$$\mathbb{E}^x \left[e^{-\beta(\tau \wedge n)} (X_{\tau \wedge n} - K) \right] = x - K \mathbb{E}^x \left[e^{-\beta(\tau \wedge n)} \right] \leq x. \quad (2.13)$$

Letting $n \rightarrow \infty$, and from Fatou's lemma, we get for any $\tau \in \mathcal{T}$

$$\mathbb{E}^x \left[e^{-\beta \tau} g(X_\tau) \right] \leq \mathbb{E}^x \left[e^{-\beta \tau} X_\tau \right] = \mathbb{E}^x \left[\lim_{n \rightarrow \infty} e^{-\beta(\tau \wedge n)} X_{\tau \wedge n} \right] \leq \liminf_{n \rightarrow \infty} \mathbb{E}^x \left[e^{-\beta(\tau \wedge n)} X_{\tau \wedge n} \right] = x$$

and so

$$w(x) \leq x.$$

Conversely, from (2.13), by taking $\tau = n$, we have

$$w(x) \geq x - Ke^{-\beta n}$$

and so by sending n to infinity, we obtain

$$w(x) \geq x.$$

Therefore, $w(x) = x$ as announced. And, it's never optimal to stop in the case $\beta = \mu$.

2.2 Under finite stopping times

Denote by \mathcal{T}_F the set of all finite stopping times (i.e. for any $\tau \in \mathcal{T}_F$, $\mathbb{P}(\tau < \infty) = 1$). In this section, we look for optimal stopping times in \mathcal{T}_F instead of searching over all stopping times in \mathcal{T} . Then, our objective function is given by

$$u(x) = \sup_{\tau \in \mathcal{T}_F} \mathbb{E}^x [e^{-\beta\tau} g(X_\tau)] \quad (2.14)$$

where X and g are defined by (1.2) and (1.3) respectively, and $\beta > \mu$ is the discounting factor. Analogously, we have the following theorem

Theorem 2. *Suppose the underlying process is a geometric Brownian motion defined in (1.2) and let v be given by (2.8). Then v solves (2.14) and hence is the price of a perpetual American call option g defined in (1.3) ; if $\mathbb{P}(\tau_{L^*} < \infty) = 1$, the stopping time τ_L defined in (2.1) for $L = L^*$ is optimal.*

Proof. (a) First we want to show that v is an upper bound for (2.14). Since u is the value function derived from a smaller set of stopping times than w , clearly, we have

$$v(x) = w(x) \geq u(x).$$

(b). Next we want to show that v is a lower bound for (2.14). It's similar to part (b) of proof for theorem 1. The only difference lies in the last step of (2.12). We have

$$v(x) = \mathbb{E}^x \left[e^{-\beta\tau_{L^*}} g(X_{\tau_{L^*}}) \right] = \mathbb{E}^x \left[\lim_{n \rightarrow \infty} e^{-\beta(\tau_{L^*} \wedge n)} g(X_{\tau_{L^*} \wedge n}) \right].$$

Using Fatou's lemma, we obtain

$$\mathbb{E}^x \left[\lim_{n \rightarrow \infty} e^{-\beta(\tau_{L^*} \wedge n)} g(X_{\tau_{L^*} \wedge n}) \right] \leq \liminf_{n \rightarrow \infty} \mathbb{E}^x \left[e^{-\beta(\tau_{L^*} \wedge n)} g(X_{\tau_{L^*} \wedge n}) \right].$$

Since $\tau_{L^*} \wedge n$ is a finite stopping time, we derive

$$\liminf_{n \rightarrow \infty} \mathbb{E}^x \left[e^{-\beta(\tau_{L^*} \wedge n)} g(X_{\tau_{L^*} \wedge n}) \right] \leq \sup_{\tau \in \mathcal{T}_F} \mathbb{E}^x \left[e^{-\beta\tau} g(X_\tau) \right].$$

Therefore, we have

$$v(x) \leq \sup_{\tau \in \mathcal{T}_F} \mathbb{E}^x \left[e^{-\beta\tau} g(X_\tau) \right] = u(x).$$

Chapter 3

Cautious Stochastic Choice, Optimal Stopping and Deliberate Randomisation

3.1 Introduction

It is well recognized that individual decision making is not fully captured by expected utility theory and many non-expected utility theories have been developed with the aim of providing a better fit to observed behaviour. Many of these alternative theories have been well studied in a static setting, but recently there has been much interest in studying non-expected utility preferences in dynamic settings which describe timing problems arising in real world decisions. Examples of theoretical work in this vein include Ebert and Strack (2015) and in experimental settings, Oprea et al (2009). This section considers agents who face optimal timing decisions in a dynamic setting and who exhibit cautious stochastic choice (CSC).

Cerreia-Vioglio et al (2015), (2017) (see also Maccheroni (2002)) develop a theory of cautious stochastic choice in a static decision making setting. The agent aims to select a best lottery from a given set. Under CSC the agent has a family of possible utility functions in mind. For a given lottery, and for each utility, the agent computes the certainty equivalent. The agent then values the lottery via the worst-case certainty equivalent. Finally the agent chooses the best lottery which maximizes this value. Since CSC does not satisfy the quasi-convexity property (See Chapter 4, (4.2) for definition), agents may benefit from mixing (see Cerreia-Vioglio et al (2017)). Our goal is to ask if these results from a static setting also apply in an optimal stopping problem. We want to understand if CSC agents also seek

to randomise in a dynamic setting. In this section we consider a continuous time optimal stopping model for the sale of an asset in which the price process is given by a one-dimensional time-homogeneous diffusion. If the agent were an expected utility maximizer, it is well known that the optimal stopping rule is given by the first exit time of the price process from an interval, ie. a pure threshold strategy. We formulate an optimal stopping problem with CSC as follows. The agent has a family of utility functions and for a given stopping rule (in an appropriate class of admissible strategies), for each utility, computes the certainty equivalent. The worst-case is then taken over utilities. The goal is to find the stopping rule which maximizes the worst-case certainty equivalent value.

By considering a stylized but tractable example, we can show the optimal strategy is not necessarily of threshold form. For this example we can calculate the optimal stopping rule and show that it is a non-trivial mixture of threshold strategies. We then consider two realistic models where the asset price follows exponential Brownian motion. In the first model the family of utilities are *S*-shaped and reference level dependent. The second model uses a family of concave utility functions. These examples show CSC agents do randomise in realistic, dynamic optimal stopping settings.

3.2 Literature

A consistent finding in experimental studies of individual decision making is the phenomenon of stochastic or random choice. When subjects are asked to choose from the same set of options many times, they are inconsistent in their choices. Patterns of stochastic choice were first recorded by Tversky (1969) and many studies have replicated, explored and extended his results (see Agranov and Ortoleva (2017) for recent findings and a comprehensive overview, and amongst others Dwenger et al (2013), Hey and Orme (1994)). In particular, recent studies of Agranov and Ortoleva (2017) and Dwenger et al (2013) interpret their experimental results as suggesting the main force is a *deliberate* desire of participants to randomise. Much of this evidence is gathered in static settings. Recently, researchers have studied *dynamic* settings which can better reflect the real decision making situations individuals face in economics and finance (eg. Oprea et al (2009)). Strack and Viefers (2017) conduct an experiment in a sophisticated asset selling task whereby subjects played sixty-five rounds during which they could sell their stock. In each round they observe a path of the market price which follows a random walk with positive expected return. Strack and Viefers (2017) present evidence that players do not play cut-off or threshold

strategies over gains - they do not behave time-consistently within rounds 75% of the time, and visit the same price level three times on average before stopping at it. In contrast to the behaviour of an EU agent, our CSC agent does not only use pure threshold strategies and instead prefers mixed or randomised strategies, consistent with this body of recent evidence. The CSC agent is deliberately randomizing, which is again, in line with the recent experimental findings as described above.

There is a large body of theoretical models which were developed to capture the phenomenon of stochastic or random choice. CSC falls into the class of stochastic models postulating that stochasticity is a *deliberate choice* of the agent.¹ Deliberate randomisation (Machina (1985)) emerges in non-EU settings such as prospect theory (see Wakker (2010) in a static setting, and Henderson, Hobson and Tse (2017) and He et al (2017) in dynamic setups). There are fewer models capturing the phenomenon of stochastic choice in the dynamic setting of a stopping problem. Strack and Viefers (2017) combine random utility with regret preferences in a stopping context. Henderson, Hobson and Tse (2017) and He et al (2017) show randomised strategies are optimal in a stopping model with prospect theory preferences. The largest class of stopping models are the bounded rationality Drift Diffusion models (DDM) of which the work of Fudenberg, Strack and Strzalecki (2017) is a recent example. Our CSC model contributes a new dynamic optimal stopping model to this class of stochastic choice models in the literature.

3.3 CSC model in a static setting (Cerreia-Vioglio et al (2017))

In this section, we present the CSC model in a static setting and give a mixing result of Cerreia-Vioglio et al (2017).

We first establish notation and review the theory for the optimal liquidation of an asset in the classical setting of a maximizer of expected utility. For J an interval subset of \mathbb{R} , let F_{\uparrow}^J be the set of increasing functions $F_{\uparrow}^J = \{f : J \mapsto \mathbb{R}; f \text{ increasing}\}$. For $f \in F_{\uparrow}^J$ we can define the left-continuous inverse f^{-1} . For K a subset of \mathbb{R}^d let $\mathcal{P}(K)$ be the set of Borel probability measures on K .

Consider an interval $I \subseteq \mathbb{R}$ of possible monetary prizes. Let $\Delta = \{\nu : \nu \in \mathcal{P}(I)\}$ be the set of lotteries over I and let Q be a subset of Δ . The agent has a set of

¹There are two other main classes of models of stochastic choice. In *random utility* models, subjects maximize a well defined utility function but this changes stochastically over time (eg. Gul and Pesendorfer (2006)). In models of *bounded rationality*, agents have well defined and stable preferences but may not make the best choice because of bounded rationality (see Johnson and Ratcliff (2013) for a review).

utility functions $\mathcal{W} \subseteq F_{\uparrow}^I$. Given a lottery $q \in Q$ and a utility $u \in \mathcal{W}$, we denote by $\mathbb{E}^q(u)$ the expected utility of u with respect to q , that is $\mathbb{E}^q(u) = \int_I u(x)q(dx)$. The certainty equivalent of q with respect to u is defined as

$$C_q^u = u^{-1}(\mathbb{E}^q(u)). \quad (3.1)$$

Under the CSC paradigm of Cerreia-Vioglio et al (2015, 2017) (see also earlier work of Maccheroni (2002)) the agent chooses a best lottery from Q by displaying cautious behaviour: *the evaluation for any given lottery $q \in Q$ is determined by $V^q = \min_{u \in \mathcal{W}} C_q^u$; the optimal strategy is to choose the lottery $\bar{q} \in Q$ which maximizes V^q .* This involves both minimization and maximization steps. Note that typically I , Q and \mathcal{W} are taken to be compact so that the optimizers exist.

Now we want to allow the agent to mix over lotteries. Let $co(Q)$ denote the convex hull of Q . Then $\rho = \rho^\lambda \in co(Q)$, represents a compound lottery obtained through a randomisation $\lambda \in \mathcal{P}(Q)$. If λ is a discrete distribution over $q \in Q$ we have $\rho^\lambda = \sum \lambda_i q_i$; more generally, $\rho^\lambda = \int_Q \lambda(dq)q$ is a measure on I given by $\rho^\lambda(dx) = \int_Q \lambda(dq)q(dx)$. For a lottery $\rho^\lambda \in co(Q)$ we can define the expected utility of u with respect to ρ^λ by $\mathbb{E}^{\rho^\lambda}(u) = \int_I u(x)\rho^\lambda(dx) = \int_Q \lambda(dq)\mathbb{E}^q(u)$, and then the certainty equivalent of ρ^λ with respect to u is $C_{\rho^\lambda}^u = u^{-1}(\mathbb{E}^{\rho^\lambda}(u))$. Let $V^{\rho^\lambda} = \min_{u \in \mathcal{W}} C_{\rho^\lambda}^u$. Then an optimal randomised lottery is given by

$$\rho^* = \rho^*(Q) \in \arg \max_{\rho^\lambda \in co(Q)} V^{\rho^\lambda}.$$

In this static setting, Cerreia-Vioglio et al (2017) show that mixing over two lotteries may improve the worst case certainty equivalent.

Suppose $Q = \{p, q\}$ and $\mathcal{W} = \{u, v\}$. If for an arbitrary constant $K > 0$, we have

$$\begin{aligned} C_p^u &> K > C_p^v \\ C_q^u &< K < C_q^v \end{aligned}$$

then, a linear combination of p and q is better than any one of them in terms of smallest certainty equivalent. To see this, note that for $\lambda \in (0, 1)$ and $\rho = \rho^\lambda = \lambda p + (1 - \lambda)q$

$$\mathbb{E}^\rho(u) = \lambda \mathbb{E}^p(u) + (1 - \lambda) \mathbb{E}^q(u) = \lambda u(C_p^u) + (1 - \lambda) u(C_q^u) > \lambda u(K) + (1 - \lambda) u(C_q^u) > u(C_q^u)$$

and since u is strictly increasing, it follows that

$$u^{-1}(\mathbb{E}^\rho(u)) > u^{-1}(u(C_q^u)) = C_q^u.$$

A similar argument gives that $v^{-1}(\mathbb{E}^\rho(v)) > C_p^v$. Then

$$\min_{w \in \mathcal{W}} w^{-1} [\mathbb{E}^\rho(w)] > \max\{C_p^v, C_q^u\} = \max_{r \in Q} \min_{w \in \mathcal{W}} C_r^w.$$

It follows that in a static setting it can be optimal to take a mixed strategy.

3.4 The Optimal stopping models

Let $\mathcal{L}(Z)$ denote the law of a random variable Z . If $Y = (Y_t)_{t \geq 0}$ is a stochastic process and \mathcal{S} is a class of stopping times then let $Q^Y(\mathcal{S}) = \{\mathcal{L}(Y_\tau); \tau \in \mathcal{S}\}$. Let δ_y be the point mass at y .

We work on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$. Let $Y = (Y_t)_{t \geq 0}$ be a (\mathbb{F}, \mathbb{P}) -stochastic process on this probability space. Let I^Y be the state space of Y and let \bar{I}^Y be the closure of I^Y . We suppose that Y is a regular, time-homogeneous diffusion with initial value $Y_0 = y$ which lies in the interior of I^Y . Further we suppose that $\lim_{t \uparrow \infty} Y_t$ exists. A sufficient condition for this is Assumption 1 below. Y represents the price process of the asset.

3.4.1 Optimal stopping under expected utility

Let U be an increasing utility function, $U \in F_{\uparrow}^{\bar{I}^Y}$. For a maximizer of expected utility the objective is to find the certainty equivalent

$$C^{EU}(\mathcal{S}) = \sup_{\tau \in \mathcal{S}} U^{-1}(\mathbb{E}^y[U(Y_\tau)]) = U^{-1} \left(\sup_{\tau \in \mathcal{S}} \mathbb{E}^y[U(Y_\tau)] \right) \quad (3.2)$$

over a suitable class \mathcal{S} of stopping times. We introduce three classes of stopping times

- \mathcal{T} , the class of all stopping times;
- \mathcal{T}_T , the class of (pure) threshold stopping times;
- \mathcal{T}_R , the class of randomised threshold stopping times.

The set of pure threshold stopping times includes stopping immediately and can be written as

$$\mathcal{T}_T = \cup_{(\beta, \gamma) \in \mathcal{D}} \{\tau_{\beta, \gamma}^Y\}, \quad (3.3)$$

where $\tau_{\beta, \gamma}^Y = \inf_{u \geq 0} \{u : Y_u \notin (\beta, \gamma)\}$ and the union is taken over (β, γ) in an appropriate set $\mathcal{D}^Y \subseteq ([-\infty, y] \cap \bar{I}^Y) \times ([y, \infty] \cap \bar{I}^Y)$ which we describe below.

In order to be able to define the set of randomised threshold stopping times \mathcal{T}_R we suppose that \mathcal{F}_0 is rich enough as to support any probability measure η on \mathcal{D} , and that the dynamics of Y are independent of η . Then we define a randomised stopping time τ_η by

$$\tau_\eta = \inf_{u \geq 0} \{u : Y_u \notin (\Theta_\beta, \Theta_\gamma) \text{ where } \Theta = (\Theta_\beta, \Theta_\gamma) \text{ is } \mathcal{F}_0 \text{ measurable and has law } \eta\}$$

and set

$$\mathcal{T}_R = \{\tau_\eta; \eta \in \mathcal{P}(\mathcal{D})\}. \quad (3.4)$$

Note that $\mathcal{T}_T \subset \mathcal{T}_R \subset \mathcal{T}$.

Often, the best way to solve (3.2) is via a change of scale. Let s be a strictly increasing function such that $X = s(Y)$ is a local martingale. Such a function s exists under very mild conditions on Y (see Rogers and Williams (2000)), and is called a scale function. For example, if Y solves the SDE $dY_t = \sigma(Y_t)dB_t + \mu(Y_t)dt$ then $s = s(z)$ is a solution to $\frac{1}{2}\sigma(z)^2 s'' + \mu(z)s' = 0$. Note that if s is a scale function then so is any affine transformation of s and so we may choose any convenient normalization for s . Then $U(Y_\tau) = g(X_\tau)$ where $g = U \circ s^{-1}$ and (3.2) can be rewritten as

$$C^{EU}(\mathcal{S}) = \sup_{\tau \in \mathcal{S}} U^{-1}(\mathbb{E}^x[g(X_\tau)]) = s^{-1} \left(\sup_{\tau \in \mathcal{S}} g^{-1}(\mathbb{E}^x[g(X_\tau)]) \right) \quad (3.5)$$

where $x = s(y)$. Since the scale function s is fixed, in finding the optimal stopping rule it is sufficient to consider $\sup_{\tau \in \mathcal{S}} g^{-1}(\mathbb{E}^x[g(X_\tau)])$. The state space of X is $I^X = s(I^Y)$. Then $\bar{I}^X = s(\bar{I}^Y)$. If I^X is not bounded below then for any level γ in the interior of I^X with $\gamma \geq x$ the first hitting time $H_\gamma^X = \inf_{u \geq 0} \{u : X_u = \gamma\}$ is finite almost surely and $C^{EU}(\mathcal{T}) = \sup_{\gamma \in I^X} U^{-1}g(\gamma) = \sup\{\gamma : \gamma \in I^Y\} = \max\{\gamma : \gamma \in \bar{I}^Y\}$. We want to exclude this degenerate case. Hence we make the following assumption:

Assumption 1. $I^X = s(I^Y)$ is bounded below. Then, without loss of generality we may assume that the lower limit of I^X is zero. Any accessible boundary point for X is absorbing.

The upper limit of I^X may be finite or infinite. Note that since X is a non-negative local martingale $\lim_{t \uparrow \infty} X_t$ exists and hence $\lim_{t \uparrow \infty} Y_t$ exists. We do not exclude τ such that $\mathbb{P}(\tau = \infty) > 0$ and on the set $\tau = \infty$ we define $X_\tau = \lim_{t \uparrow \infty} X_t$.

This is why we want to consider \bar{I}^X as well as I^X . Then \mathcal{T} is the set of all stopping times, and not just finite stopping times.

Example 1. Suppose Y is geometric Brownian motion: $dY_t = \sigma Y_t dB_t + \mu Y_t dt$, $Y_0 = y$. Let $\psi = 1 - \frac{2\mu}{\sigma^2}$. Y has state space $I^Y = (0, \infty)$. Provided $\psi \neq 0$ we have $s(z) = \text{sgn}(\psi)z^\psi$. (If $\psi = 0$ then $s(z) = \ln z$ is the scale function.) If $\psi \leq 0$ then $s(0) = -\infty$. This is equivalent to $2\mu \geq \sigma^2$, in which case Y hits arbitrarily high price levels with probability one and the optimal stopping problem is degenerate. If $\psi > 0$ then $I^X = (0, \infty)$. For $\psi > 0$, $\lim_{t \rightarrow \infty} X_t = 0 \in \bar{I}^X \setminus I^X$. Moreover, fix $y > 0$, and for an arbitrary constant $b > y$, we have

$$\mathbb{P}_y(\tau_b < \infty) = \lim_{n \uparrow \infty} \mathbb{P}_y(\tau_b < \tau_{1/n}) = \lim_{n \uparrow \infty} \frac{s(y) - s(1/n)}{s(b) - s(1/n)} = \left(\frac{y}{b}\right)^\psi < 1$$

Therefore, geometric Brownian motion can produce non-trivial examples.

Remark 3. Under Assumption 1 the process X is a non-negative local martingale, and hence a supermartingale. Further, for any stopping time $\mathbb{E}^x[X_\tau] \leq x$. If I^X is bounded above then X is a martingale and $\mathbb{E}^x[X_\tau] = x$, but for many examples $I^X = (0, \infty)$ or $[0, \infty)$ and then there exist τ for which $\mathbb{E}^x[X_\tau] < x$.

We do not make a concavity assumption on U . Monotonicity is preserved under the transformation $U \mapsto g$, but in general concavity is not. Indeed, if g is concave then typically stopping immediately ($\tau = 0$) is optimal (see Figure 3.1).

Note that $\tau_{\beta, \gamma}^Y = \inf_{u \geq 0} \{u : Y_u \notin (\beta, \gamma)\} = \inf_{u \geq 0} \{u : X_u \notin (s(\beta), s(\gamma))\} =: \tau_{s(\beta), s(\gamma)}^X$. Hence \mathcal{T}_T has the alternative representation

$$\mathcal{T}_T = \cup_{(a,b) \in \mathcal{D}^X} \{\tau_{a,b}^X\},$$

for an appropriate set \mathcal{D}^X . The right space to choose is $\mathcal{D}^X = ([0, x] \cap \bar{I}^X) \times ([x, \infty] \cap \bar{I}^X)$. Then \mathcal{D}^Y can be expressed as

$$\mathcal{D}^Y = [s^{-1}(0), y) \times [y, s^{-1}(\infty)]$$

\mathcal{T}_R can also be rewritten as $\mathcal{T}_R = \{\tau_\eta^X : \eta \in \mathcal{P}(\mathcal{D}^X)\}$ where

$$\tau_\eta^X = \inf_{u \geq 0} \{u : X_u \notin (\Theta_\beta, \Theta_\gamma) \text{ where } \Theta = (\Theta_\beta, \Theta_\gamma) \text{ is } \mathcal{F}_0 \text{ measurable and has law } \eta\}$$

Recall $Q^X(\mathcal{S})$ is the set of possible laws of the stopped X -process, over stopping times in \mathcal{S} . We have the following lemma.

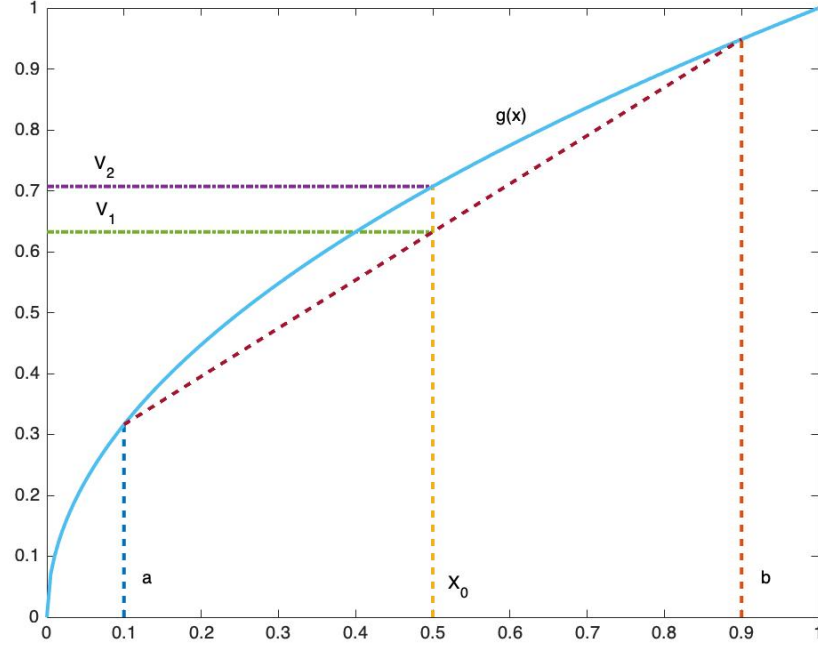


Figure 3.1: For concave function g and any stopping time $\tau_{a,b}$, the expected payoff is V_1 ; V_2 is the payoff from stopping immediately. $V_2 \geq V_1$.

- Lemma 1.** 1. If I^X is bounded then $Q^X(\mathcal{T}) = \{\mathcal{L}(X_\tau) : \tau \in \mathcal{T}\} = \{\nu \in \mathcal{P}(\bar{I}^X) : \int_{\bar{I}^X} z\nu(dz) = x\}$. If I^X is not bounded above then $Q^X(\mathcal{T}) = \{\mathcal{L}(X_\tau) : \tau \in \mathcal{T}\} = \{\nu \in \mathcal{P}(\bar{I}^X) : \int_{\bar{I}^X} z\nu(dz) \leq x\}$.
2. If I^X is bounded then $Q^X(\mathcal{T}_T) = \{\mathcal{L}(X_\tau) : \tau \in \mathcal{T}_T\} = \delta_x \cup \left(\bigcup_{\beta < x < \gamma; \beta, \gamma \in \bar{I}^X} \chi_{a,b}^x \right)$ where $\chi_{a,b}^x$ is the mixture of point masses $\chi_{a,b}^x = \frac{x-a}{b-a}\delta_b + \frac{b-x}{b-a}\delta_a$. If I^X is not bounded above then $Q^X(\mathcal{T}_T) = \{\mathcal{L}(X_\tau) : \tau \in \mathcal{T}_T\} = (\bigcup_{0 \leq \beta \leq x} \delta_\beta) \cup \left(\bigcup_{0 \leq \beta < x < \gamma < \infty} \chi_{a,b}^x \right)$.
3. In both cases $Q^X(\mathcal{T}_T) \subset Q^X(\mathcal{T}_R) = Q^X(\mathcal{T})$.

Proof. 1. Suppose I^X is unbounded. The fact that $Q^X(\mathcal{T}) \subseteq \{\nu \in \mathcal{P}(\bar{I}^X) : \int_{\bar{I}^X} z\nu(dz) \leq x\}$ follows from Remark 3. The fact that we have equality follows from the fact that by Skorokhod's Embedding Theorem any $\nu \in \mathcal{P}([0, \infty))$ with $\int z\nu(dz) \leq x$ can be obtained as the law of X_τ for a stopping time τ . See Pedersen and Peskir (2001) or Cox and Hobson (2004). The case of bounded X has a similar proof.

2. This is immediate from the definition of pure threshold rules.

3. This follows from Lemma 7, Chapter 4. The intuition is that for any $\nu \in Q^X(\mathcal{T})$ we can find a stopping time $\tau \in \mathcal{T}_R$ such that $\mathcal{L}(X_\tau) = \nu$. This type of task is known as Skorokhod embedding problem (See Skorokhod (1965)). \square

Note that the certainty equivalent depends only on the law of X_τ . The following result is classical. (In discrete time see Karni and Safra (1990), in mathematical finance see Dayanik and Karatzas (2003) and recently in discrete time in Strack and Viefers (2017). For a textbook treatment, see Chapter 4, Peskir and Shiryaev (2006).)

Proposition 1. 1. $C^{EU}(\mathcal{T}_T) = C^{EU}(\mathcal{T}_R) = C^{EU}(\mathcal{T})$.

2. $C^{EU}(\mathcal{T}) = U^{-1}(g^{cv}(s(y)))$ where g^{cv} is the smallest concave majorant of $g = U \circ s^{-1}$.

Corollary 1. In trying to find the optimal stopping rule in the classical (single utility) case it is sufficient to restrict attention to pure threshold strategies of the form $\tau = \tau_{a,b}^Y$.

Proof of Proposition 1. This proposition is standard but we provide a short proof which will have parallels to our method in the CSC case. The results will follow if we can show that $\sup_{\tau \in \mathcal{T}_T} \mathbb{E}^x[g(X_\tau)] = \sup_{\tau \in \mathcal{T}} \mathbb{E}^x[g(X_\tau)] = g^{cv}(x)$. Note that $g = U \circ s^{-1}$ is increasing. For $d \geq 0$ suppose $g(z) \leq c + dz$. Then, by the supermartingale property of X , $\mathbb{E}^x[g(X_\tau)] \leq c + d\mathbb{E}^x[X_\tau] \leq c + dx$. Taking an infimum over $c, d \geq 0$ for which $c + dz \geq g(z)$ we find $\sup_{\tau \in \mathcal{T}} \mathbb{E}^x[g(X_\tau)] = g^{cv}(x)$. Conversely, either $g(x) = g^{cv}(x)$ and then $\sup_{\tau \in \mathcal{T}_T} \mathbb{E}^x[g(X_\tau)] \geq \mathbb{E}^x[g(X_0)] = g^{cv}(x)$ or there exists a largest interval I_x with endpoints $\{a_x, b_x\}$ such that $x \in I_x$ and $g(z) < g^{cv}(z)$ on I_x . If $g(a_x) = g^{cv}(a_x)$ and $g(b_x) = g^{cv}(b_x)$ then $\sup_{\tau \in \mathcal{T}_T} \mathbb{E}^x[g(X_\tau)] \geq \mathbb{E}^x[g(X_{\tau_{a,b}^X})] = g^{cv}(x)$. Otherwise, there exist $a_n \rightarrow a_x$ and $b_n \rightarrow b_x$ such that $g(a_n) \rightarrow g^{cv}(a_x)$ and $g(b_n) \rightarrow g^{cv}(b_x)$. Then $\sup_{\tau \in \mathcal{T}_T} \mathbb{E}^x[g(X_\tau)] \geq \limsup \mathbb{E}^x[g(X_{\tau_{a_n, b_n}^X})] = g^{cv}(x)$. \square

3.4.2 Optimal stopping under Cautious Stochastic Choice

Our goal in this section is to develop an optimal stopping model with CSC. Let Y be a time-homogeneous diffusion with state space I^Y . Let $\mathcal{W}^Y \subseteq F_{\uparrow}^{I^Y}$ be a set of increasing utility functions. The goal is to find $\sup_{\tau \in \mathcal{S}} \inf_{u \in \mathcal{W}^Y} u^{-1}(\mathbb{E}[u(Y_\tau)])$, where τ is chosen from a suitable set of stopping times \mathcal{S} . We define $\Delta^Y = \mathcal{P}(\bar{I}^Y)$. Recall that $Q^Y(\mathcal{S}) = \{\nu : \nu = \mathcal{L}(Y_\tau); \tau \in \mathcal{S}\}$. By Lemma 1 we have $Q^Y(\mathcal{T}_T) \subset Q^Y(\mathcal{T}_R) = Q^Y(\mathcal{T}) \subseteq \Delta^Y$.

As in the classical, single-utility setting, it is often convenient to work with the process X in natural scale rather than Y . We set $\mathcal{W}^X = \{g = u \circ s^{-1}; u \in \mathcal{W}^Y\}$.

Define $\Delta^X = \mathcal{P}(\bar{I}^X)$. Again by Lemma 1 we have $Q^X(\mathcal{T}_T) \subset Q^X(\mathcal{T}_R) = Q^X(\mathcal{T}) \subseteq \Delta^X$.

For a fixed stopping time τ and a fixed utility $u \in \mathcal{W}$ we define the certainty equivalent

$$C_\tau^u = u^{-1}(\mathbb{E}[u(Y_\tau)]) = u^{-1}(\mathbb{E}[g(X_\tau)]) = s^{-1}(g^{-1}(\mathbb{E}[g(X_\tau)])). \quad (3.6)$$

Once we have minimized over utilities the value function for a single stopping time is $V_\tau = \inf_{u \in \mathcal{W}^Y} C_\tau^u$. Under CSC the optimal stopping problem is to find $V(\mathcal{S}) = \sup_{\tau \in \mathcal{S}} V_\tau$ where \mathcal{S} is a set of stopping times. Since V_τ depends on the stopping time only through the law of the stopped process we have

$$\begin{aligned} V(\mathcal{S}) &= \sup_{\nu \in Q^Y(\mathcal{S})} \inf_{u \in \mathcal{W}^Y} u^{-1} \left(\int u(z) \nu(dz) \right) \\ &= s^{-1} \left(\sup_{\nu \in Q^X(\mathcal{S})} \inf_{g \in \mathcal{W}^X} g^{-1} \left(\int g(z) \nu(dz) \right) \right) \end{aligned} \quad (3.7)$$

and $\tau^* \in \operatorname{argmax}_{\tau \in \mathcal{S}} V_\tau$. In particular, we want to consider $\mathcal{S} = \mathcal{T}$, $\mathcal{S} = \mathcal{T}_R$ and $\mathcal{S} = \mathcal{T}_T$.

The following result is key to solving (3.7):

Proposition 2. *For V defined in (3.7), we have $V(\mathcal{T}_T) \leq V(\mathcal{T}_R) = V(\mathcal{T})$.*

The proposition follows that Lemma 1. Lemma 1 characterizes the sets $Q^X(\mathcal{S})$ for various sets \mathcal{S} . However, the sets \mathcal{T} , \mathcal{T}_T and \mathcal{T}_R do not depend on whether we consider stopping times for the process X or Y . Hence $\{\nu : \nu \in Q^Y(\mathcal{S})\} = \{\eta_\# s : \eta \in Q^X(\mathcal{S})\}$ where, by definition $\eta_\# s(A) = \eta(s(A))$. From our perspective, the content of Proposition 2 is that $V(\mathcal{T}_R) = V(\mathcal{T})$. The first main result of this paper, is to show that, unlike in the classical case (see Proposition 1, (1), we may have $V(\mathcal{T}_T) < V(\mathcal{T}_R)$.

3.5 A stylized example

Our goal in this section is to give an example for which we can prove that the optimal stopping rule is not a pure threshold strategy. Instead there is an optimal stopping rule which is a non-trivial mixture of threshold stopping rules. The example is highly stylized, and deliberately simple, and this allows us to give a full and complete solution, i.e. we are able to solve for the optimal mixed threshold rule. Crucially, as we will see in the next section, the characteristic features are shared with some realistic, non-stylized examples.

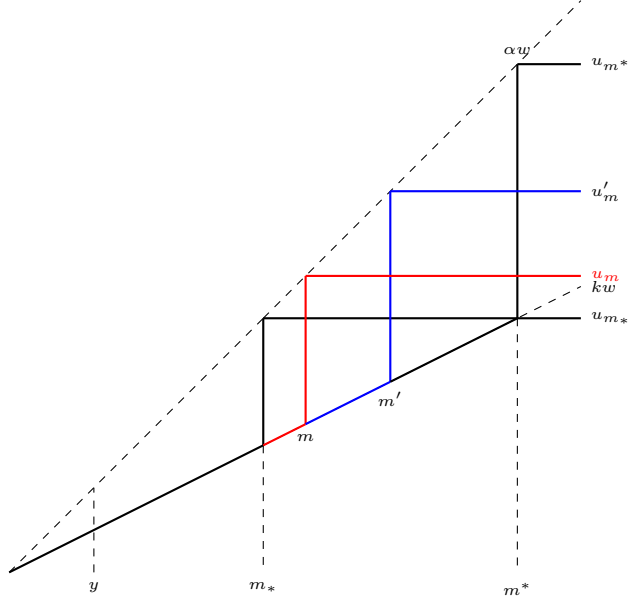


Figure 3.2: The family of utility functions $\mathcal{W} = \{u_m; m \in \mathcal{M}\}$, where u_m is defined for $m \in \mathcal{M}$ by $u_m(w) = kw, 0 \leq w < m$ and $u_m(w) = \alpha m, w \geq m$ for constants $\alpha > k > 0$. Let $\mathcal{M} = [m_*, m^*]$ with $m_* > \frac{\alpha y}{k}$ and $m^* = \frac{\alpha m_*}{k}$.

We work with a process Y which is already in natural scale, and a family of payoff functions $\{u_m\}_{m \in \mathcal{M}}$ where $\mathcal{M} \subset \mathbb{R}$. The process Y is assumed to be bounded below (without loss of generality by zero) and unbounded above, to be a local martingale and to have initial value $y > 0$. Then Y is a supermartingale. The canonical example is if Y is a Brownian motion started at y and absorbed at zero. Alternatively, we may consider Y to be geometric Brownian motion with zero drift. The goal in this section is to give an example for which

$$V(\mathcal{T}_T) < V(\mathcal{T}_R) = V(\mathcal{T}).$$

Hence, there is no pure threshold strategy which is optimal within class of all stopping rules.

Fix constants $\alpha > k > 0$, together with $m_* > \frac{\alpha y}{k}$. Set $m^* = \frac{\alpha m_*}{k}$ and let \mathcal{M} be the interval $\mathcal{M} = [m_*, m^*]$. For $m \in \mathcal{M}$ define $u_m : I \equiv [0, \infty) \mapsto [0, \infty)$ by

$$u_m(w) = \begin{cases} kw, & 0 \leq w < m; \\ \alpha m, & w \geq m; \end{cases} \quad (3.8)$$

and set $\mathcal{W} = \{u_m; m \in \mathcal{M}\}$. Figure 3.2 illustrates the family of utilities described here.

Note that the results generalize to utility functions which replace $u_m(w) = \alpha m$ with $u_m(w) = J(m)$ for $w \geq m$ in (3.8), where J is a strictly increasing function with $J(m) > km$. We will consider this more general case in Section 3.5.4.

3.5.1 Pure threshold strategies

Our first result is that in the stylized example there is no pure threshold strategy which outperforms the trivial strategy of stopping immediately. Note that since Y is a supermartingale and since $u_m(z) \leq \alpha z$, we have for all $\tau \in \mathcal{T}$

$$\mathbb{E}[u_m(Y_\tau)] \leq \alpha \mathbb{E}[Y_\tau] \leq \alpha y < km_*.$$

Since $u_m^{-1}(w) = \frac{w}{k}$ for $w < km_*$ we have for any $\tau \in \mathcal{T}$

$$u_m^{-1}(\mathbb{E}[u_m(Y_\tau)]) = \frac{1}{k} \mathbb{E}[u_m(Y_\tau)]. \quad (3.9)$$

Recall $\tau_{\beta,\gamma} = \inf\{s : Y_s \notin (\beta, \gamma)\}$. For $m \in \mathcal{M}$ and $0 \leq \beta \leq y \leq \gamma$ let $\mathcal{G}_{\beta,\gamma}^m$ be the expected utility associated with the stopping time $\tau_{\beta,\gamma}$ and the utility function u_m and let $C_{\beta,\gamma}^m$ be the certainty equivalent: we have $\mathcal{G}_{\beta,\gamma}^m = \mathbb{E}[u_m(Y_{\tau_{\beta,\gamma}})]$ and $C_{\beta,\gamma}^m = (u_m)^{-1}(\mathbb{E}[u_m(Y_{\tau_{\beta,\gamma}})]) = \frac{1}{k} \mathcal{G}_{\beta,\gamma}^m$. Then

$$\mathcal{G}_{\beta,\gamma}^m = \begin{cases} ky & \gamma \in [y, m) \\ \left(\alpha m \frac{y-\beta}{\gamma-\beta} + k\beta \frac{\gamma-y}{\gamma-\beta} \right) & \gamma \geq m \end{cases} \quad (3.10)$$

$$C_{\beta,\gamma}^m = \begin{cases} y & \gamma \in [y, m) \\ \frac{1}{k} \left(\alpha m \frac{y-\beta}{\gamma-\beta} + k\beta \frac{\gamma-y}{\gamma-\beta} \right) & \gamma \geq m. \end{cases} \quad (3.11)$$

Note that for each $m \in \mathcal{M}$, $\mathcal{G}_{\beta,\gamma}^m$ and $C_{\beta,\gamma}^m$ are non-increasing in γ for $\gamma \geq m^*$. Also, for each $m \in \mathcal{M}$, and $\gamma \leq m^*$, $\mathcal{G}_{\beta,\gamma}^m$ and $C_{\beta,\gamma}^m$ are non-increasing in β for $0 \leq \beta \leq y$.

Lemma 2. For all $0 \leq \beta \leq y \leq \gamma$, $\inf_{m \in \mathcal{M}} C_{\beta,\gamma}^m \leq y$.

Moreover, $\sup_{\beta,\gamma} \inf_{m \in \mathcal{M}} C_{\beta,\gamma}^m = y$.

Proof. If $\gamma \in [y, m_*)$ then $C_{\beta,\gamma}^m = y$ for all $m \in \mathcal{M}$.

If $\gamma \geq m^*$ then using the fact that $C_{\beta,\gamma}^m$ is increasing in m for $m \leq \gamma$ and $\alpha m_* = km^*$,

$$\begin{aligned}
\inf_{m \in \mathcal{M}} C_{\beta, \gamma}^m &= C_{\beta, \gamma}^{m_*} = \frac{1}{k} \left(\alpha m_* \frac{y - \beta}{\gamma - \beta} + k \beta \frac{\gamma - y}{\gamma - \beta} \right) \\
&= m_* \frac{y - \beta}{\gamma - \beta} + \beta \frac{\gamma - y}{\gamma - \beta} \\
&= y - \frac{(y - \beta)(\gamma - m_*)}{(\gamma - \beta)} \leq y.
\end{aligned}$$

Finally, if $\gamma \in [m_*, m^*)$ then $\inf_{m \in \mathcal{M}} C_{\beta, \gamma}^m \leq C_{\beta, \gamma}^{m^*} = y$.

The first statement of the lemma follows from consideration of the three possible cases. The second statement follows from the first, given that for all m , $C_{\beta, y}^m = y$. \square

Theorem 3. *In our stylized example no pure threshold strategies outperforms stopping immediately.*

Proof of Theorem 3. The result follows immediately from the previous lemma that in our stylized example $V(\mathcal{T}_T) = y$. From the perspective of the worst agent, any pure threshold strategy can only generate at best a certainty equivalent which is the same as the certainty equivalent from selling the asset immediately. \square

3.5.2 Improvement with randomisation between Two Upper Thresholds

The goal of this section is to show that there are mixtures of threshold strategies which outperform the best pure threshold strategies. In addition we will develop some intuition which we can use to motivate the derivation of the optimal randomised strategy.

The remarks after (3.11) suggest that it is not sensible to use upper thresholds above m^* , and that it is sufficient to only consider lower thresholds which are set to zero. (This result is proved in Lemma 4) In this section we consider using stopping rules which are a mixture of $\tau_{0, \gamma}$ and $\tau_{0, \epsilon}$ for $m_* \leq \gamma < \epsilon \leq m^*$. If τ is this mixed stopping rule and $\epsilon < m^*$ then $u_{m^*}(Y_\tau) = kY_\tau$ and the certainty equivalent is equal to y . So, if we hope to outperform pure threshold rules we must take $\epsilon = m^*$.

Let \mathcal{T}_2^0 be the set of stopping rules obtained from mixing two pure threshold strategies, both with lower threshold 0, and one with upper threshold at m^* , and the other with upper threshold in $[m_*, m^*)$. Then $\mathcal{T}_2^0 = \{\tau_\gamma^\theta : \theta \in [0, 1], \gamma \in [m_*, m^*)\}$ where $\tau_\gamma^\theta = \tau_{0, \gamma}$ with probability θ and $\tau_\gamma^\theta = \tau_{0, m^*}$ otherwise. The randomisation over $\tau_{0, \gamma}$ and τ_{0, m^*} takes place at $t = 0$. Set $\mathcal{C}_\gamma^{m, \theta} = u_m^{-1} \left(\mathbb{E}[u_m(Y_{\tau_\gamma^\theta})] \right)$. Then, by

the linearity of u_m^{-1} over the relevant domain (recall (3.9))

$$\begin{aligned}\mathcal{C}_\gamma^{m,\theta} &= u_m^{-1} \left(\theta \mathbb{E}[u_m(Y_{\tau_0,\gamma})] + (1-\theta) \mathbb{E}[u_m(Y_{\tau_0,m^*})] \right) \\ &= \frac{\theta}{k} \mathbb{E}[u_m(Y_{\tau_0,\gamma})] + \frac{1-\theta}{k} \mathbb{E}[u_m(Y_{\tau_0,m^*})] \\ &= \theta C_{0,\gamma}^m + (1-\theta) C_{0,m^*}^m.\end{aligned}$$

It follows that

$$\mathcal{C}_\gamma^{m,\theta} = \begin{cases} y \left[\theta + (1-\theta) \frac{\alpha m}{k m^*} \right] & m_* \leq \gamma < m \leq m^* \\ \frac{y \alpha m}{k} \left[\frac{\theta}{\gamma} + \frac{(1-\theta)}{m^*} \right] & m_* \leq m \leq \gamma \leq m^*. \end{cases}$$

Fix $\gamma \in [m_*, m^*)$ and θ . As a function of m , $H_\gamma^\theta(m) = \mathcal{C}_\gamma^{m,\theta}$ is increasing in m on both $[m_*, \gamma]$ and $(\gamma, m^*]$ with a jump down at γ . It follows that (with $\mathcal{C}_\gamma^{\gamma+,\theta} = \lim_{m \downarrow \gamma} \mathcal{C}_\gamma^{m,\theta}$)

$$\begin{aligned}\inf_m \mathcal{C}_\gamma^{m,\theta} &= \min\{\mathcal{C}_\gamma^{m_*,\theta}, \mathcal{C}_\gamma^{\gamma+,\theta}\} \\ &= y \min \left\{ \frac{\theta m^*}{\gamma} + (1-\theta); \theta + (1-\theta) \frac{\alpha \gamma}{k m^*} \right\}.\end{aligned}\tag{3.12}$$

Continuing to fix γ but allowing the mixture parameter θ to vary, the first term in the minimum in (3.12) is increasing in θ whereas the second is decreasing in θ . Also, at $\theta = 0$, $\mathcal{C}_\gamma^{m_*,0} = y < \frac{\alpha \gamma}{k m^*} y = \mathcal{C}_\gamma^{\gamma+,0}$ and at $\theta = 1$, $\mathcal{C}_\gamma^{m^*,1} = \frac{m^*}{\gamma} y > y = \mathcal{C}_\gamma^{\gamma+,1}$. It follows that $\inf_m \mathcal{C}_\gamma^{m,\theta}$ is maximized over θ at the value of θ for which $\mathcal{C}_\gamma^{m_*,\theta} = \mathcal{C}_\gamma^{\gamma+,\theta}$, namely $\theta = \theta^*(\gamma)$ where

$$\theta^*(\gamma) = \frac{\gamma - m_*}{\frac{m_* m^*}{\gamma} + \gamma - 2m_*} \in (0, 1).$$

Then

$$\max_{\theta \in [0,1]} \inf_m \mathcal{C}_\gamma^{m,\theta} = y \frac{m^* - m_*}{\frac{m_* m^*}{\gamma} + \gamma - 2m_*}.$$

Finally, we find the maximizer over γ is $\gamma = \gamma^*$ where $\gamma^* = \sqrt{m_* m^*}$, and then $\theta^*(\gamma^*) = \frac{1}{2}$ and

$$\max_{\gamma \in \mathcal{M}} \max_{\theta \in [0,1]} \inf_m \mathcal{C}_\gamma^{m,\theta} = \mathcal{C}_{\gamma^*}^{m_*,\theta^*(\gamma^*)} = y \left[\frac{1 + \sqrt{\frac{m^*}{m_*}}}{2} \right] > y.$$

Hence, $V(\mathcal{T}_2^0) > V(\mathcal{T}_T)$ and *a fortiori* $V(\mathcal{T}) > V(\mathcal{T}_T)$.

Theorem 4. *In our stylized example the best strategy outperforms the best pure threshold strategy.*

In addition to the above result, we can learn something from our analysis about the optimal mixture of thresholds. First we expect that there must be a positive probability that we take an upper threshold of m^* , else the certainty equivalent associated with u_{m^*} is y . Second, by considering the problem for finite mixtures of upper thresholds, we expect that the certainty equivalent associated with u_m should be constant over m .

3.5.3 The Optimal Solution

Next we record some useful properties about $\mathcal{G}_{\beta,\gamma}^m$ $\mathcal{C}_{\beta,\gamma}^m$ which follow immediately from the definitions in (3.11).

Lemma 3. 1. *For each $m \in \mathcal{M}$, $\mathcal{G}_{\beta,\gamma}^m$ and $\mathcal{C}_{\beta,\gamma}^m$ are non-increasing in γ for $\gamma \geq m^*$. Hence, for $\gamma \geq m^*$, $\inf_m \mathcal{C}_{\beta,\gamma}^m \leq \inf_m \mathcal{C}_{\beta,m^*}^m$.*

2. *For each $m \in \mathcal{M}$, and $\gamma \leq m^*$, $\mathcal{G}_{\beta,\gamma}^m$ and $\mathcal{C}_{\beta,m}^m$ are non-increasing in β for $0 \leq \beta \leq y$. Hence for $\gamma \leq m^*$, $\inf_m \mathcal{C}_{\beta,\gamma}^m \leq \inf_m \mathcal{C}_{0,\gamma}^m$.*

Consider the spaces of probability measures $\mathcal{P}(\mathcal{M})$, $\mathcal{P}(\{0\} \times \mathcal{M})$ and $\mathcal{P}([0, y] \times [y, \infty])$. For any $A \times B \in \mathcal{P}(\{0\} \times \mathcal{M})$, we have A is a point mass at 0 of size 1. Thus, there is an obvious 1-1 correspondence between measures $\tilde{\zeta} \in \mathcal{P}(\mathcal{M})$ and $\hat{\zeta} \in \mathcal{P}(\{0\} \times \mathcal{M})$ given by $\hat{\zeta}(\{0\} \times B) = \tilde{\zeta}(B)$. Write $\tilde{\eta} = p(\hat{\eta})$ for this correspondence. Similarly, we can define a map $P : \mathcal{P}([0, y] \times [y, \infty]) \mapsto \mathcal{P}(\{0\} \times \mathcal{M})$ by $P(\eta) = \hat{\eta}$ where

$$\hat{\eta}(\{0\}, d\gamma) = \begin{cases} \hat{\eta}(\{0\}, \delta_{m_*}) = \int_{\beta} \int_{\gamma \leq m_*} \eta(d\beta, d\gamma) & , \gamma \leq m_* \\ \hat{\eta}(\{0\}, d\gamma) = \int_{\beta} \eta(d\beta, d\gamma) & , \gamma \in (m_*, m^*) \\ \hat{\eta}(\{0\}, \delta_{m^*}) = \int_{\beta} \int_{\gamma \geq m^*} \eta(d\beta, d\gamma) & , \gamma \geq m^* \end{cases}$$

Recall $\mathcal{G}_{\beta,\gamma}^m = \mathbb{E}[u_m(Y_{\tau_{\beta,\gamma}})]$ and $\mathcal{C}_{\beta,\gamma}^m = u_m^{-1}(\mathcal{G}_{\beta,\gamma}^m)$. Let η be a probability measure on $[0, y] \times [y, \infty]$. We can define a randomised stopping time $\tau = \tau_{\eta}$ by generating a random variable $\Theta = (\Theta_{\beta}, \Theta_{\gamma})$ with law η on $[0, y] \times [y, \infty]$ and setting $\tau = \tau_{\Theta_{\beta}, \Theta_{\gamma}}$. Then we define \mathcal{G}_{η}^m to be the expected utility from using the randomised stopping rule τ_{η} :

$$\mathcal{G}_{\eta}^m = \int_{[0,y] \times [y,\infty]} \eta(d\beta, d\gamma) \mathcal{G}_{\beta,\gamma}^m = \mathbb{E}^{\eta}[u_m(Y_{\tau_{\beta,\gamma}})]$$

Finally we set \mathcal{C}_{η}^m to be the certainty equivalent: $\mathcal{C}_{\eta}^m = u_m^{-1}(\mathcal{G}_{\eta}^m)$.

Lemma 4. $\mathcal{G}_\eta^m \leq \mathcal{G}_{P(\eta)}^m$.

Proof. From Lemma 3, for $m \in \mathcal{M}$ and $\gamma \geq m^*$, $\mathcal{G}_{\beta,\gamma}^m \leq \mathcal{G}_{\beta,m^*}^m \leq \mathcal{G}_{0,m^*}^m$. Similarly, for $m \in \mathcal{M}$ and $\gamma < m_*$, $\mathcal{G}_{\beta,\gamma}^m = y = \mathcal{G}_{\beta,m_*}^m = \mathcal{G}_{0,m_*}^m$. Then defining $M(\gamma) = (\gamma \vee m_*) \wedge m^*$ we have $\mathcal{G}_{\beta,\gamma}^m \leq \mathcal{G}_{0,M(\gamma)}^m$. It follows that for all $(\beta, \gamma) \in ([0, y) \times [y, \infty])$ we have $\mathcal{G}_{\beta,\gamma}^m \leq \mathcal{G}_{0,M(\gamma)}^m$. Then for any $\eta \in \mathcal{P}([0, y) \times [y, \infty])$, writing $\hat{\eta} = P(\eta)$,

$$\mathcal{G}_\eta^m = \int \eta(d\beta, d\gamma) \mathcal{G}_{\beta,\gamma}^m \leq \int \hat{\eta}(d\beta, d\gamma) \mathcal{G}_{\beta,\gamma}^m = \mathcal{G}_{\hat{\eta}}^m.$$

□

For $\tilde{\zeta} \in \mathcal{P}(\mathcal{M})$ define $\mathcal{C}_{\tilde{\zeta}}^m = \mathcal{C}_{\hat{\zeta}}^m$ where $\hat{\zeta} = p^{-1}(\tilde{\zeta})$.

Corollary 2.

$$\sup_{\eta \in \mathcal{P}([0, y) \times [y, \infty])} \inf_{m \in \mathcal{M}} \mathcal{C}_\eta^m = \sup_{\zeta \in \mathcal{P}(\{0\} \times \mathcal{M})} \inf_{m \in \mathcal{M}} \mathcal{C}_\zeta^m = \sup_{\tilde{\zeta} \in \mathcal{P}(\mathcal{M})} \inf_{m \in \mathcal{M}} \mathcal{C}_{\tilde{\zeta}}^m$$

Proof. We have

$$\begin{aligned} & \sup_{\eta \in \mathcal{P}([0, y) \times [y, \infty])} \inf_{m \in \mathcal{M}} u_m^{-1} \left(\int \eta(d\beta, d\gamma) \mathcal{G}_{\beta,\gamma}^m \right) \\ & \leq \sup_{\hat{\eta} \in \mathcal{P}(\{0\} \times \mathcal{M})} \inf_{m \in \mathcal{M}} u_m^{-1} \left(\int \hat{\eta}(d\beta, d\gamma) \mathcal{G}_{\beta,\gamma}^m \right) \\ & = \sup_{\tilde{\eta} \in \mathcal{P}(\mathcal{M})} \inf_{m \in \mathcal{M}} u_m^{-1} \left(\int \tilde{\eta}(d\gamma) \mathcal{G}_{0,\gamma}^m \right) \end{aligned} \quad (3.13)$$

Since $\mathcal{P}(\{0\} \times \mathcal{M}) \subseteq \mathcal{P}([0, y) \times [y, \infty])$ the inequality (3.13) is an equality, and the result follows. □

Let \mathcal{T}_R^0 be the subset of \mathcal{T}_R such that the lower threshold in the randomisation mixture is always at zero, and the upper threshold is in \mathcal{M} . Then

$$\mathcal{T}_R^0 = \{\tau_\eta : \eta \in \mathcal{P}(\{0\} \times \mathcal{M})\}. \quad (3.14)$$

Proposition 3. $V(\mathcal{T}_R^0) = V(\mathcal{T}_R) = V(\mathcal{T})$.

Proof of Proposition 3. Let ζ be a probability measure on $\{0\} \times \mathcal{M} \subseteq [0, y) \times [y, \infty)$. We can identify ζ with a probability measure $\tilde{\zeta}$ on \mathcal{M} and then $\mathcal{G}_{\tilde{\zeta}}^m = \int_{\mathcal{M}} \tilde{\zeta}(d\gamma) \mathcal{G}_{0,\gamma}^m$ and $\mathcal{C}_{\tilde{\zeta}}^m = u_m^{-1}(\mathcal{G}_{\tilde{\zeta}}^m)$.

Corollary 2 shows that $V(\mathcal{T}_R^0) = V(\mathcal{T}_R)$ and it is sufficient to only consider threshold strategies in the mixture with lower bound at 0 and upper bound in \mathcal{M} . The fact that $V(\mathcal{T}_R) = V(\mathcal{T})$ follows from Proposition 2 and $Q(\mathcal{T}_R) = Q(\mathcal{T})$. \square

Thus, in the stylized example and when considering optimal mixtures of threshold strategies it is sufficient to restrict attention to mixtures in which the lower threshold is always zero and the upper threshold is contained in \mathcal{M} . Our calculation of the optimal strategy in the CSC setting is based on the following general proposition. Let \mathcal{Z} be a set and let \mathcal{N} be a measurable space. Let $D : \mathcal{Z} \times \mathcal{N} \mapsto \mathbb{R}$ be a map and set $D_*(z) = \inf_{n \in \mathcal{N}} D(z, n)$ and $D^* = D^*(\mathcal{Z}) = \sup_{z \in \mathcal{Z}} D_*(z)$.

Proposition 4. *Suppose there exist $\mathcal{Z}_0 \subseteq \mathcal{Z}$, $\mathcal{N}_0 \subseteq \mathcal{N}$, $z^* \in \mathcal{Z}_0$, $\nu \in \mathcal{P}(\mathcal{N}_0)$, a family $(h_n)_{n \in \mathcal{N}_0}$ of strictly increasing functions $h_n : \mathbb{R} \mapsto \mathbb{R}$ and constants \hat{D} , \hat{H} such that*

$$\begin{aligned} D(z^*, n) &\geq \hat{D} \text{ on } \mathcal{N} \text{ with } D(z^*, n) = \hat{D} \text{ on } \mathcal{N}_0 \\ \int_{\mathcal{N}_0} \nu(dn) h_n(D(z, n)) &\leq \hat{H} \text{ on } \mathcal{Z} \text{ with } \int_{\mathcal{N}_0} \nu(dn) h_n(D(z, n)) = \hat{H} \text{ on } \mathcal{Z}_0 \end{aligned}$$

Then, for any $z \in \mathcal{Z}$

$$D_*(z) \leq D_*(z^*) = D^*$$

In our interpretation we take \mathcal{Z} to be either the space of stopping rules or the space of attainable laws or the set of randomisations of the levels of lower and upper thresholds. (Since our problem is law invariant, the final result will be equivalent.) \mathcal{Z}_0 is a space of relevant stopping rules or attainable laws or randomisations, for example the set of randomised threshold rules for which the upper barrier lies in some interval. \mathcal{N} is a parameterization of the space of utility functions and \mathcal{N}_0 is a set of relevant utility functions. We may have $\mathcal{N}_0 \neq \mathcal{N}$ if there are utility functions for which the certainty equivalent is never the lowest over the family of utility functions. See Section 3.5.4 for an example. Then $D(z, n)$ is the certainty equivalent using utility function u_n and stopping rule z ; $D_*(z)$ is the CSC value of the stopping rule z .

The first idea behind the proof is that we expect the certainty equivalent value of the optimal stopping rule to be constant across the set of (relevant) utility functions. If not, we might expect to be able to improve the certainty equivalent value under the worst utility, at the expense of the certainty equivalent values of those utilities which have a higher certainty equivalent value. This would raise the CSC value. Hence we expect $D(z^*, n)$ is constant on \mathcal{N}_0 for the optimal choice z^* .

The second idea is that we want there to be only one (randomised threshold) stopping rule for which the certainty equivalent is constant (across all relevant

utilities). This possibility is achieved by a requirement that no stopping rule can achieve a certainty equivalent value which exceeds that of another relevant stopping rule, uniformly across all relevant utilities.

Proof. Take $z \in \mathcal{Z}$ and w in \mathcal{Z}_0 . Suppose $D(z, n) > D(w, n)$ for all $n \in \mathcal{N}_0$. Then $h_n(D(z, n)) > h_n(D(w, n))$ for all $n \in \mathcal{N}_0$ and $\hat{H} \geq \int_{\mathcal{N}_0} \nu(dn) h_n(D(z, n)) > \int_{\mathcal{N}_0} \nu(dn) h_n(D(w, n)) = \hat{H}$ contradicting the hypothesis of the theorem. Hence, for any $z \in \mathcal{Z}, w \in \mathcal{Z}_0$ there exists a non-empty set $\mathcal{N}_{z,w} \subseteq \mathcal{N}_0$ such that $D(z, n) \leq D(w, n)$ on $\mathcal{N}_{z,w}$. Taking $w = z^*$, we have

$$D(z, n) \leq D(z^*, n) \text{ for } n \in \mathcal{N}_{z,z^*}$$

It follows that

$$D_*(z) = \inf_{n \in \mathcal{N}} D(z, n) \leq \inf_{n \in \mathcal{N}_{z,z^*}} D(z, n) \leq \inf_{n \in \mathcal{N}_{z,z^*}} D(z^*, n)$$

By assumption, we have

$$\inf_{n \in \mathcal{N}_{z,z^*}} D(z^*, n) = \hat{D} = \inf_{n \in \mathcal{N}} D(z, n) = D_*(z^*) = D^*.$$

□

Theorem 5. Suppose $\tilde{\eta} \in \mathcal{P}(\mathcal{M})$ is a mixture of a point mass at m^* of size θ^* and an absolutely continuous measure ρ on (m_*, m^*) with density $C^* \gamma^{-\frac{\alpha}{\alpha-k}}$ where

$$\theta^* = \frac{1}{\left(\frac{\alpha}{k}\right)^{\frac{\alpha}{\alpha-k}} - \frac{\alpha-k}{k}} \quad C^* = \frac{\frac{\alpha}{\alpha-k} (m^*)^{\frac{k}{\alpha-k}}}{\left(\frac{\alpha}{k}\right)^{\frac{\alpha}{\alpha-k}} - \frac{\alpha-k}{k}}$$

Then an optimal strategy is to take a randomised strategy with mixture distribution $\hat{\eta}$ where $\hat{\eta}(\{0\}, d\gamma) = \tilde{\eta}(d\gamma)$. The corresponding value function is

$$V = y \frac{\left(\frac{\alpha}{k}\right)^{\frac{k}{\alpha-k}}}{\left(\frac{\alpha}{k}\right)^{\frac{k}{\alpha-k}} - \frac{\alpha-k}{\alpha}}.$$

Proof of Theorem 5. The idea is to apply Proposition 4. To this end take $\mathcal{Z}_0 = \mathcal{Z} = \mathcal{P}(\mathcal{M})$, $\mathcal{N}_0 = \mathcal{N} = \mathcal{M}$ and set

$$f(\gamma, m) = \mathbb{E}[g_m(Y_{\tau_0, \gamma})] = \mathcal{G}_{0, \gamma}^m = \begin{cases} ky & \gamma \in [m_*, m) \\ \frac{\alpha m y}{\gamma} & \gamma \in [m, m^*]. \end{cases}$$

Note that for $\chi \in \mathcal{P}(\mathcal{M})$, $\mathcal{C}_\chi^m = g_m^{-1}(\int_{\mathcal{M}} \mathcal{G}_{0,\gamma}^m \chi(d\gamma))$. Take $h_m = g_m$ and $D(\chi, m) = \mathcal{C}_\chi^m$. Then

$$\int \nu(dm) h_m(\mathcal{C}_\zeta^m) = \int \nu(dm) \mathcal{G}_\zeta^m = \int_{\mathcal{M}} \nu(dm) \int_{\mathcal{M}} f(\gamma, m) \zeta(d\gamma) = \int_{\mathcal{M}} \zeta(d\gamma) \int_{\mathcal{M}} f(\gamma, m) \nu(dm).$$

Then by Proposition 4, if we can find $\tilde{\zeta} \in \mathcal{P}(\mathcal{M})$ such that $\mathcal{C}_{\tilde{\zeta}}^m$ does not depend on m and ν such that $\int_{\mathcal{M}} f(\gamma, m) \nu(dm)$ does not depend on γ then $\tilde{\zeta}$ characterizes the optimal mixture of thresholds.

The required conditions follow from the next two lemmas. \square

Lemma 5. For $\tilde{\eta}$ as in the statement of Theorem 5, $\mathcal{C}_{\tilde{\eta}}^m = \frac{1}{k} \int_{\mathcal{M}} f(\gamma, m) \tilde{\eta}(d\gamma)$ is independent of m .

Proof. It follows from the definition of C^* and θ^* that

$$1 = \int_{\mathcal{M}} C^* \gamma^{-\frac{\alpha}{\alpha-k}} d\gamma + \theta^*$$

so that $\tilde{\eta}$ is a probability measure on \mathcal{M} . Then

$$\begin{aligned} \frac{1}{ky} \int_{\mathcal{M}} \mathcal{G}_\gamma^m \tilde{\eta}(d\gamma) &= \int_{[m_*, m)} C^* \gamma^{-\frac{\alpha}{\alpha-k}} d\gamma + \int_{[m, m^*)} \frac{\alpha m}{k} C^* \gamma^{-1-\frac{\alpha}{\alpha-k}} d\gamma + \frac{\theta^* \alpha m}{k m^*} \\ &= -\frac{\alpha-k}{k} C^* \gamma^{-\frac{k}{\alpha-k}} \Big|_{m_*}^m - \frac{\alpha-k}{\alpha} \frac{\alpha m}{k} C^* \gamma^{-\frac{\alpha}{\alpha-k}} \Big|_m^{m^*} + \frac{\theta^* \alpha m}{k m^*} \\ &= \frac{\alpha-k}{k} C^* m_*^{-\frac{k}{\alpha-k}} - \frac{\alpha-k}{k} C^* m(m^*)^{-\frac{\alpha}{\alpha-k}} + \frac{\theta^* \alpha m}{k m^*} \end{aligned}$$

This does not depend on m since $\theta^* = \frac{\alpha-k}{\alpha} C^* (m^*)^{1-\frac{\alpha}{\alpha-k}}$. \square

Lemma 6. Let $\lambda = \frac{\alpha}{k} > 1$. Let ν be a mixture of an atom of size $\phi = (\lambda^{\frac{\lambda}{\lambda-1}} - \lambda + 1)^{-1}$ at m_* and an absolutely continuous measure ζ on \mathcal{M} with density $Dm^{\frac{2-\lambda}{\lambda-1}}$ where $D = \frac{\lambda}{\lambda-1} \phi m_*^{-\frac{1}{\lambda-1}}$. Then $\int_{\mathcal{M}} f(\gamma, m) \nu(dm)$ does not depend on γ .

Proof. Set $\beta = \frac{2-\lambda}{\lambda-1}$. Then $\beta + 1 = \frac{1}{\lambda-1}$ and $\beta + 2 = \frac{\lambda}{\lambda-1}$. With the absolutely continuous component of ν having density Dm^β we have

$$\begin{aligned} \frac{1}{y} \int f(\gamma, m) \nu(dm) &= \int_{[m_*, \gamma]} \frac{\alpha m}{k\gamma} \nu(dm) + \int_{(\gamma, m^*]} \nu(dm) \\ &= \frac{\alpha m_*}{k\gamma} \phi + \frac{\alpha}{k\gamma} \int_{(m_*, \gamma]} Dm^{\beta+1} dm + \int_{(\gamma, m^*]} Dm^\beta dm \\ &= \frac{\alpha m_*}{k\gamma} \phi + \frac{\alpha}{k\gamma} \frac{Dm^{\beta+2}}{\beta+2} \Big|_{m_*}^\gamma + \frac{Dm^{\beta+1}}{\beta+1} \Big|_\gamma^{m^*} \\ &= \frac{m^*}{\gamma} \left[\phi - \frac{Dm_*^{\beta+1}}{(\beta+2)} \right] + \gamma^{\beta+1} D \left[\frac{\alpha}{k(\beta+2)} - \frac{1}{\beta+1} \right] + \frac{D(m^*)^{\beta+1}}{(\beta+1)} \end{aligned}$$

The two square brackets in this last expression are zero by the choice of D and β . Hence $\int f(\gamma, m)\nu(dm)$ does not depend on γ . \square

It is worth highlighting here that the optimal stopping rule is not unique and although in Theorem 5 we find the optimal mixed threshold rule, there are other stopping times which are also optimal. In other words, suppose $\tau \in \mathcal{T}_R$ is a randomised threshold rule (which is not a pure threshold rule): then there exist other stopping times $\tau' \in \mathcal{T}$ for which $\mathcal{L}(X_\tau) = \mathcal{L}(X_{\tau'})$ or equivalently $\mathcal{L}(Y_\tau) = \mathcal{L}(Y_{\tau'})$.

3.5.4 A Generalized Example

Fix $L > 0$ and suppose $R \in (L, \infty)$. Let $J : [L, R] \mapsto \mathbb{R}$ be a continuously differentiable function with $J(z) > z$ and such that $\sup_{z \in [L, R]} \frac{J(z)}{z} < \kappa < \infty$, where $\kappa > 1$ is constant. Let $K : [L, R] \mapsto \mathbb{R}$ be the largest increasing differentiable function such that $K \leq J$ and $\int_L^R d\beta \left| \frac{K'(\beta)}{K(\beta) - \beta} \right| < \infty$. Suppose $J(L) = K(L) \geq R$ and that the set $\{x : K(x) = J(x)\}$ is the union of a finite set of intervals. We write $\{x : K(x) = J(x)\} = \cup_{i=1}^N [\ell_i, r_i]$. Then $\ell_1 = L$ and $r_N = R$.

Let $\mathcal{A} = [L, R]$ and for $\alpha \in \mathcal{A}$ define $u_\alpha : [0, \infty) \mapsto \mathbb{R}$ by

$$u_\alpha(z) = \begin{cases} z & 0 \leq z < \alpha; \\ J(\alpha) & z \geq \alpha. \end{cases} \quad (3.15)$$

Let $\mathcal{W} = \{u_\alpha : \alpha \in [L, R]\}$.

Let Y be Brownian motion started at $y \in (0, \frac{L}{\kappa})$, and absorbed at 0. Consider the problem of finding

$$\sup_{\tau} \inf_{u \in \mathcal{W}} u^{-1}(\mathbb{E}[u(Y_\tau)]).$$

Note that for any $\alpha \in [L, R]$, we have $u_\alpha(z) \leq \kappa z$. It follows that for any stopping time τ , $\mathbb{E}[u_\alpha(Y_\tau)] \leq \kappa \mathbb{E}[Y_\tau] \leq \kappa y < L$. But $u^{-1}(x) = x$ over this range. Hence the u^{-1} may be omitted in the definition of the Cautious Stochastic Utility in this example.

Theorem 6. *Let θ be given by*

$$\theta = \left[1 + \frac{1}{R} \int_L^R d\alpha \frac{\alpha K'(\alpha)}{K(\alpha) - \alpha} \exp \left(\int_\alpha^R d\beta \frac{K'(\beta)}{K(\beta) - \beta} \right) \right]^{-1}$$

and let $\rho : [L, R] \mapsto \mathbb{R}_+$ be given by

$$\rho(\alpha) = \frac{\theta}{R} \left\{ \frac{\alpha K'(\alpha)}{K(\alpha) - \alpha} \exp \left(\int_\alpha^R d\beta \frac{K'(\beta)}{K(\beta) - \beta} \right) \right\}. \quad (3.16)$$

Let $\tilde{\eta} \in \mathcal{P}([L, R])$ be the probability measure with density ρ on $[L, R]$ and a point mass of size θ at R .

Then an optimal strategy is to take a randomized threshold strategy with mixture distribution $\hat{\eta}$ where $\hat{\eta}(\{0\}, d\gamma) = \tilde{\eta}(d\gamma)$ and $\hat{\eta}$ does not charge $(0, x) \times [x, \infty]$.

Proof of Theorem 6. We apply Proposition 4. Let \mathcal{Z} be the set of probability measures on $[0, y) \times [y, \infty]$ for some $y \in (0, \frac{L}{\kappa})$ and let $\mathcal{Z}_0 \subseteq \mathcal{Z}$ be the set of probability measures with support $\{0\} \times [L, R]$. Let $\mathcal{N} = [L, R]$ and let $\mathcal{N}_0 = \{\alpha : K(\alpha) = J(\alpha)\} \subseteq \mathcal{N}$.

Then \mathcal{Z} is the set of candidate randomizations, and \mathcal{Z}_0 is a set of relevant randomizations which are not dominated by some other randomization. \mathcal{N} is a parameterization of the utility functions, and \mathcal{N}_0 is a set of utility functions such that no member dominates any other element of \mathcal{N} .

Recall the definitions of θ , ρ , $\tilde{\eta}$ and $\hat{\eta}$ from the theorem. By the choice of θ , $\tilde{\eta}$ is a probability measure on $[L, R]$. Define $\Delta : [L, R] \mapsto \mathbb{R}$ by

$$\Delta(\alpha) = \frac{\theta}{R} \exp \left(\int_{\alpha}^R d\beta \frac{K'(\beta)}{K(\beta) - \beta} \right).$$

Then Δ is differentiable and from the definition of ρ in (3.16)

$$\Delta'(\alpha) = -\frac{\Delta(\alpha)K'(\alpha)}{(K(\alpha) - \alpha)} = -\frac{\rho(\alpha)}{\alpha}. \quad (3.17)$$

Then, since $\Delta(R) = \frac{\theta}{R}$,

$$\Delta(\alpha) = \frac{\theta}{R} + \int_{(\alpha, R)} \frac{\rho(\beta)}{\beta} d\beta = \int_{(\alpha, R]} \frac{\tilde{\eta}(d\beta)}{\beta}.$$

For $\zeta \in \mathcal{P}([0, y) \times [y, \infty])$ and $\alpha \in \mathcal{N}$ define

$$D(\zeta, \alpha) = \frac{1}{y} u_{\alpha}^{-1}(\mathbb{E}[u_{\alpha}(Y_{\tau_{\zeta}})]) = \frac{1}{y} \mathbb{E}[u_{\alpha}(Y_{\tau_{\zeta}})] = \int \int \zeta(d\beta, d\gamma) \frac{u_{\alpha}(\beta)(\gamma - y) + u_{\alpha}(\gamma)(y - \beta)}{y(\gamma - \beta)}.$$

For $\hat{\zeta} \in \mathcal{P}(\{0\} \times [L, R])$, this reduces to

$$D(\hat{\zeta}, \alpha) = \int \tilde{\zeta}(d\gamma) \frac{u_{\alpha}(\gamma)}{\gamma}$$

where $\tilde{\zeta} = p(\hat{\zeta})$.

Then, for $\hat{\eta} \in \mathcal{Z}$ as in the statement of the Theorem,

$$\begin{aligned}
\frac{1}{y} \mathbb{E}[u_\alpha(Y_{\tau_{\hat{\eta}}})] &= \tilde{\eta}([L, \alpha)) + J(\alpha) \int_{(\alpha, R]} \frac{\tilde{\eta}(d\beta)}{\beta} \\
&\geq \tilde{\eta}([L, \alpha)) + K(\alpha)\Delta(\alpha) \\
&= \theta + K(L)\Delta(L) + \int_{[L, \alpha)} [\rho(\beta) + \Delta(\beta)K'(\beta) + K(\beta)\Delta'(\beta)] d\beta \\
&= \theta + J(L)\Delta(L)
\end{aligned}$$

where we use the first inequality in (3.17) to show that the integrand in the penultimate line is zero.

Then, if $\hat{D} = \theta + J(L)\Delta(L)$ we have for all $\alpha \in \mathcal{N}$,

$$\mathbb{E}[u_\alpha(Y_{\tau_{\hat{\eta}}})] \geq \hat{D}$$

with equality if $\alpha \in \mathcal{N}_0$.

It remains to show that there exists a measure ν with support in $[L, R]$ and \hat{H} such that $\int_{\mathcal{N}_0} \nu(d\alpha)D(\zeta, \alpha) = \hat{H}$ for $\zeta \in \mathcal{Z}_0$, and $\int_{\mathcal{N}_0} \nu(d\alpha)D(\zeta, \alpha) \leq \hat{H}$ for general $\zeta \in \mathcal{Z}$. (We take $h_\alpha(d) = d$ for all $\alpha \in \mathcal{N}$.)

Recall that $\{z : K(z) = J(z)\} = \cup_{i=1}^N [\ell_i, r_i]$. Let ν be the measure on $\{z : K(z) = J(z)\}$ such that ν has atoms of size ϕ_i at ℓ_i for $i = 1, 2, \dots, N$, together with a density ζ on $\cup_{i=1}^N (\ell_i, r_i)$ given by

$$\zeta(w) = \zeta_i \exp \left(\int_{\ell_i}^w \frac{(2 - J'(z))}{(J(z) - z)} dz \right).$$

Here $\phi_1 = 1$ and $\zeta_1 = \frac{\phi_1 J(L)}{L(J(L) - L)}$, and then, proceeding inductively, for $1 \leq i \leq N-1$,

$$\phi_{i+1} = \frac{(\ell_{i+1} - r_i)}{r_i(K(\ell_{i+1}) - \ell_{i+1})} \int_{\alpha \leq r_i} K(\alpha) \nu(d\alpha) \quad (3.18)$$

$$\zeta_{i+1} = \frac{1}{\ell_{i+1}(K(\ell_{i+1}) - \ell_{i+1})} \int_{\alpha \leq \ell_{i+1}} K(\alpha) \nu(d\alpha) \quad (3.19)$$

For any $\tilde{\zeta}$ with support in $[L, R]$ we can define $\hat{\zeta} = p^{-1}(\tilde{\zeta})$. Then

$$\int_{\mathcal{N}_0} \nu(d\alpha)D(\hat{\zeta}, \alpha) = \int_{\mathcal{N}_0} \nu(d\alpha) \int_{[L, R]} \frac{u_\alpha(w)}{w} \tilde{\zeta}(dw) = \int_{[L, R]} \tilde{\zeta}(dw) \left[\int_{\alpha \leq w} \nu(d\alpha) \frac{K(\alpha)}{w} + \int_{\alpha > w} \nu(d\alpha) \right].$$

First we show that $\Gamma(w) := \int_{\alpha \leq w} \nu(d\alpha) \frac{K(\alpha)}{w} + \int_{\alpha > w} \nu(d\alpha)$ is constant for $w \in \mathcal{N}_0$.

For $w \in (\ell_i, r_i)$

$$\Gamma'(w) = \zeta(w) \left[\frac{K(w)}{w} - 1 \right] - \frac{1}{w^2} \int_{\alpha \leq w} \nu(d\alpha) K(\alpha).$$

We get that Γ is constant on (ℓ_i, r_i) provided $\Upsilon(w) = 0$ where $\Upsilon(w) = \int_{\alpha \leq w} K(\alpha) \nu(d\alpha) - \zeta(w)w(K(w) - w)$. But $\Upsilon'(w) = \zeta(w)[w(2 - K'(w))] - \zeta'(w)w(K(w) - w) = 0$ by the definition of ζ . Moreover, by the definition of ζ_i in (3.19) $\Upsilon(\ell_i) = 0$. Hence $\Upsilon \equiv 0$ on $[\ell_i, r_i]$ and $\Gamma(w)$ is constant on this interval.

To prove that Γ is constant on \mathcal{N}_0 it remains only to show that $\Gamma(r_i) = \Gamma(\ell_{i+1})$. We have

$$\begin{aligned} \Gamma(\ell_{i+1}) - \Gamma(r_i) &= \frac{1}{\ell_{i+1}} \left[\int_{\alpha \leq r_i} \nu(d\alpha) K(\alpha) + K(\ell_{i+1})\phi_{i+1} \right] - \frac{1}{r_i} \left[\int_{\alpha \leq r_i} \nu(d\alpha) K(\alpha) \right] - \phi_{i+1} \\ &= \phi_{i+1} \left[\frac{K(\ell_{i+1})}{\ell_{i+1}} - 1 \right] - \left(\frac{1}{r_i} - \frac{1}{\ell_{i+1}} \right) \int_{\alpha \leq r_i} \nu(d\alpha) K(\alpha) = 0 \end{aligned}$$

by the definition of ϕ_{i+1} in (3.18).

Finally, we consider general $\eta \in \mathcal{P}([0, y] \times [y, \infty])$. Recall the definition of $\hat{\zeta} = P(\eta)$. From Lemma 4, $\mathcal{G}_\eta^m \leq \mathcal{G}_{P(\eta)}^m$ for all $m \in [L, M]$. Then $D(\eta, m) \leq D(P(\eta), m)$ for all x and

$$\int_{\mathcal{N}_0} \nu(dm) D(\eta, m) \leq \int_{\mathcal{N}_0} \nu(dm) D(P(\eta), m) = \hat{H}.$$

□

Note that if J is not strictly increasing then we have that K is constant over intervals and $\tilde{\mu}$ does not charge such intervals. The reason for this is that the corresponding u_α strictly dominate other u_β and are never the worst case utilities. For this reason they are not relevant in the CSC formulation. In the proof in the appendix, the utility functions are divided into two classes. For elements of the first class, the certainty equivalent is never smallest, and these utilities do not affect the CSC value. However, all elements of the second class are important, and we find the optimal strategy by making sure that the certainty equivalent is constant across utilities in this class, at least for the optimal mixed threshold stopping rule.

3.6 Two realistic models

In the previous section we studied a stylized liquidation problem and showed that in the CSC paradigm it is possible that the optimal strategy is not of threshold form. In this section we give two examples of more realistic models which are based on either S -shaped reference dependent utilities or on concave utilities.

3.6.1 An example with S -shaped reference dependent utilities

The first model is based on combining CSC with S -shaped reference-dependent preferences (Tversky and Kahneman (1992), see Kyle et al (2006), Henderson (2012), Ingersoll and Jin (2013), and Magnani (2016) in context of optimal stopping). This example can be shown to reduce to a form which is very closely related to the stylized example.

Suppose Y follows geometric Brownian motion and solves $dY_t = \sigma Y_t dB_t + \mu Y_t dt$ subject to $Y_0 = y$. We assume $0 < \mu < \frac{1}{2}\sigma^2$. Let $\mathcal{W}^Y = \{u_i : 1 \leq i \leq N\}$ be a family of S -shaped reference dependent utility functions with

$$u_i(z) = \begin{cases} (z - R_i)^{\delta_i} & z \geq R_i \\ -\kappa_i(R_i - z)^{\delta_i} & z < R_i \end{cases} \quad (3.20)$$

where $\{(\delta_i, R_i, \kappa_i)\}_{1 \leq i \leq N}$ is a family of parameters. Here, for each i , $\delta_i \in (0, 1)$ is a risk aversion/risk seeking parameter, $R_i > 0$ is the reference level and $\kappa_i \geq 1$ is the loss aversion parameter. Our problem is to find the CSC value

$$\sup_{\tau} \min_i u_i^{-1}(\mathbb{E}[u_i(Y_{\tau})]). \quad (3.21)$$

Define $\psi = 1 - \frac{2\mu}{\sigma^2} \in (0, 1)$ and set $s(z) = z^{\psi}$. Set $X = s(Y)$ and $x = s(y)$. Then X solves $dX_t = \psi \sigma X_t dB_t$ subject to $X_0 = x := y^{\psi} > 0$. We have X is a non-negative martingale for $t \in [0, \infty)$. Set $g_i = u_i \circ s^{-1}$ so that

$$g_i(w) = \begin{cases} (w^{1/\psi} - R_i)^{\delta_i} & w \geq R_i^{\psi} \\ -\kappa_i(R_i - w^{1/\psi})^{\delta_i} & w < R_i^{\psi} \end{cases} \quad (3.22)$$

and set $\mathcal{W}^X = \{g_i : 1 \leq i \leq N\}$. By an immediate extension of the arguments leading to (3.5) we have

$$\sup_{\tau} \min_i u_i^{-1}(\mathbb{E}[u_i(Y_{\tau})]) = s^{-1} \left(\sup_{\tau} \min_i g_i^{-1}(\mathbb{E}[g_i(X_{\tau})]) \right)$$

and hence in the search for the optimal stopping rule it is sufficient to consider the problem in natural scale for X and \mathcal{W}^X .

Families of functions \mathcal{W}^Y and \mathcal{W}^X are given in Figure 3.3 for the parameters:

$$\psi = 0.5, N = 3 \text{ and } \{(\delta_i, R_i, \kappa_i)\} = \{(0.15, 1, 2), (0.1, 2, 2), (0.08, 3, 2)\}. \quad (3.23)$$

Note that certainty equivalents are invariant under affine transformations of the

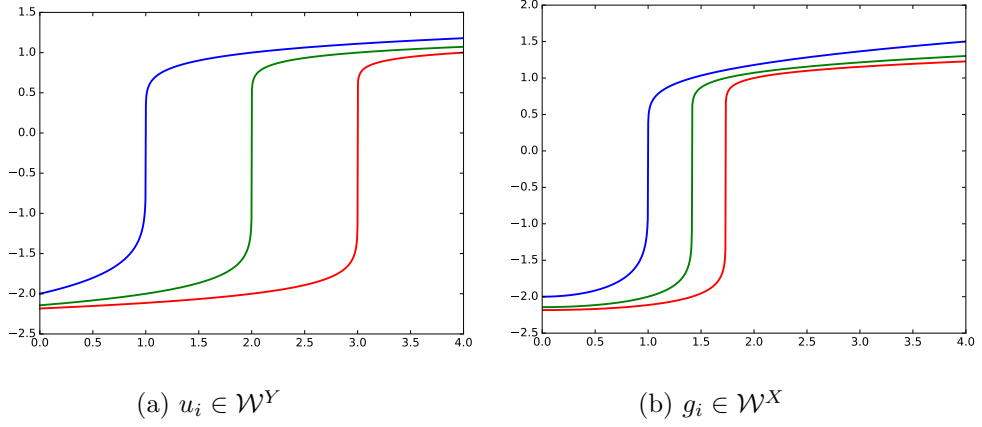


Figure 3.3: The families of S -shaped reference dependent utility functions $\mathcal{W}^Y = \{u_i : 1 \leq i \leq N\}$ with u_i defined in (3.20) and in natural scale $\mathcal{W}^X = \{g_i : 1 \leq i \leq N\}$ with g_i given in (3.22). Parameters used are $\psi = 1/2$ for the price process, $N = 3$ and $\{(\delta_i, R_i, \kappa_i)\} = \{(0.15, 1, 2), (0.1, 2, 2), (0.08, 3, 2)\}$ for the utility functions where for each i , $\delta_i \in (0, 1)$ is a risk aversion/risk seeking parameter, $R_i > 0$ is the reference level and $\kappa_i \geq 1$ is the loss aversion parameter.

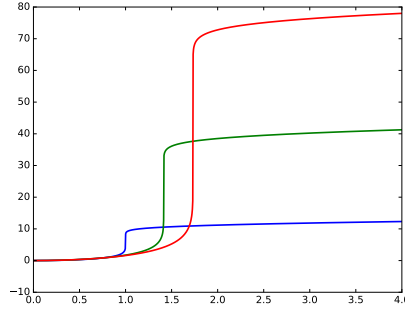


Figure 3.4: The family of transformed utility functions $\widetilde{\mathcal{W}}^X = \{\tilde{g}_i : 1 \leq i \leq N\}$ where \tilde{g}_i is given by (3.24) with $\hat{x} = 0.8$. Parameters used are $\psi = 1/2$ for the price process, $N = 3$ and $\{(\delta_i, R_i, \kappa_i)\} = \{(0.15, 1, 2), (0.1, 2, 2), (0.08, 3, 2)\}$ for the utility functions where for each i , $\delta_i \in (0, 1)$ is a risk aversion/risk seeking parameter, $R_i > 0$ is the reference level and $\kappa_i \geq 1$ is the loss aversion parameter.

objective function: if $h_{a,b}(w) = ah(w) + b$ with $a > 0$ then $h_{a,b}^{-1}(\mathbb{E}[h_{a,b}(Z)]) = h^{-1}(\mathbb{E}[h(Z)])$. Hence, without loss of generality we may replace $\mathcal{W}^X = \{g_i : 1 \leq i \leq N\}$ with $\widetilde{\mathcal{W}}^X = \{\tilde{g}_i : 1 \leq i \leq N\}$ where for fixed $\hat{x} > 0$

$$\tilde{g}_i(w) = \frac{g_i(w) - g_i(0)}{g_i(\hat{x}) - g_i(0)} \quad (3.24)$$

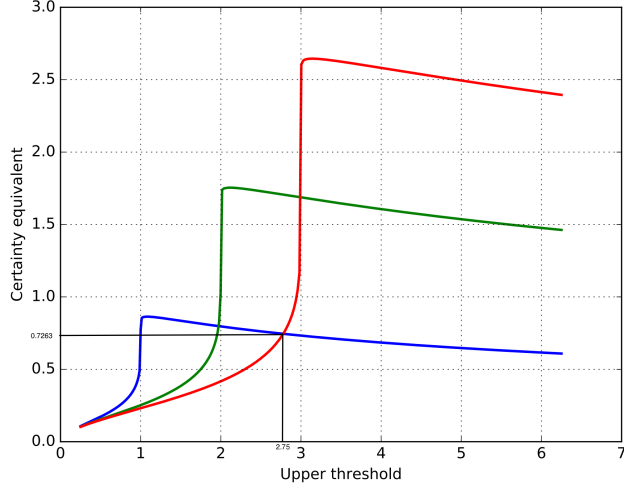


Figure 3.5: The certainty equivalent value under a pure threshold strategy $\tau_{0,\gamma}^X = \inf\{t : X_t \notin (0, \gamma)\}$ as a function of upper threshold γ for $\gamma > X_0 = x = 0.2$. The family of S -shaped utility functions u_i as defined in (3.20) are used. The best pure threshold strategy uses an upper threshold of about 2.75 and gives a CSC certainty equivalent of 0.7263, as marked on the figure. Parameters used are $\psi = 1/2$ for the price process, $N = 3$ and $\{(\delta_i, R_i, \kappa_i)\} = \{(0.15, 1, 2), (0.1, 2, 2), (0.08, 3, 2)\}$ for the utility functions where for each i , $\delta_i \in (0, 1)$ is a risk aversion/risk seeking parameter, $R_i > 0$ is the reference level and $\kappa_i \geq 1$ is the loss aversion parameter.

These linear transformations have been designed so that $\tilde{g}_i(0) = 0$ and $\tilde{g}_i(\hat{x}) = 1$ for all i . Then, the functions \tilde{g}_i are of comparable sizes over the region $[0, \hat{x}]$ and we expect that over the relevant range g_i^{-1} does not depend greatly on i . The transformed family of functions $\tilde{\mathcal{W}}^X$ are plotted in Figure 3.4. The key observation is that the resulting objective functions are similar to those studied in the stylized example in Section 3.5. Hence we expect a similar conclusion: it is not optimal to use a pure threshold rule, and instead there is an optimal stopping rule which is a non-trivial mixture of threshold rules.

Consider first the certainty equivalent from using a pure threshold strategy $\tau_{0,\gamma}^X = \inf\{t : X_t \notin (0, \gamma)\}$ for $\gamma > x = 0.2$. The certainty equivalents associated with the utilities $(u_i)_{i=1,2,3}$ as a function of the upper threshold are plotted in Figure 3.5. We see from the figure that the best pure threshold strategy uses an upper threshold of

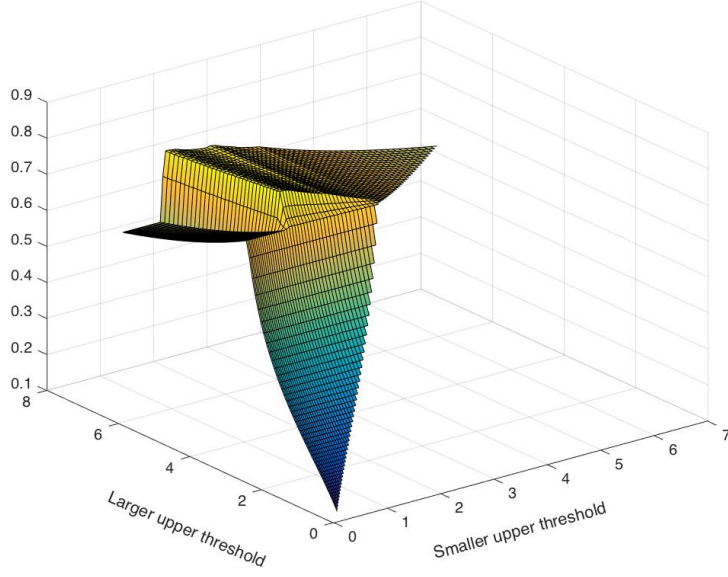


Figure 3.6: CSC value using the optimal mixture for a given pair of upper threshold levels where $X_0 = x = 0.2$. The family of S -shaped utility functions u_i as defined in (3.20) are used. The best pair of upper thresholds is 1.1, 3.1 giving a CSC certainty equivalent of 0.8368. Parameters used are $\psi = 1/2$ for the price process, $N = 3$ and $\{(\delta_i, R_i, \kappa_i)\} = \{(0.15, 1, 2), (0.1, 2, 2), (0.08, 3, 2)\}$ for the utility functions where for each i , $\delta_i \in (0, 1)$ is a risk aversion/risk seeking parameter, $R_i > 0$ is the reference level and $\kappa_i \geq 1$ is the loss aversion parameter.

approximately 2.75 and yields a CSC certainty equivalent of 0.7263.

Now suppose we are allowed to search for the best mixed threshold strategy based on two upper thresholds (with the lower threshold set to zero). Figure 3.6 shows the highest CSC value (as the mixture parameter varies) for a given pair of upper thresholds. Figure 3.7 shows how much probability mass is assigned to the smaller of the two upper thresholds.

The best strategy is to assign probability mass 0.75, 0.25 to thresholds 1.1, 3.1 respectively, giving a CSC value of 0.8368. From Figure 3.7 we see that for other pairs of thresholds, it is optimal to place all the weight on a single threshold, but for the optimal pair of thresholds the optimal strategy is a proper mixture. It follows that the best randomised strategy is strictly better than any pure threshold strategy.

We can also consider a mixture which involves at most three upper thresholds. We find that in this restricted class, the optimal randomised strategy assigns probability mass 0.76, 0.11, 0.13 to thresholds 1.1, 2.1, 3.1 respectively and gives a CSC value of 0.8425. Again, we see an improvement as we allow for mixtures over a larger number of thresholds. However, the benefit from adding more upper thresholds is

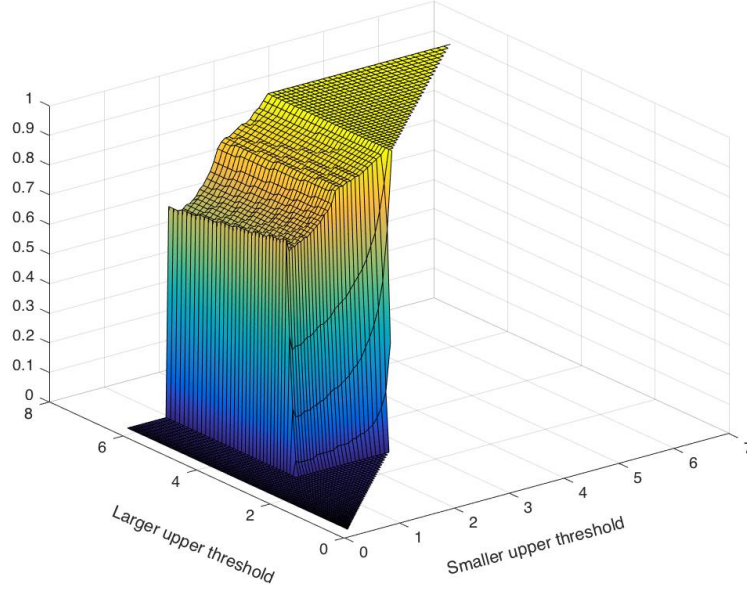


Figure 3.7: Optimal mixture distribution: the weight placed on the smaller of the upper thresholds for a given pair of upper thresholds. When both upper thresholds are large, it is optimal to not use a mixture, and only stop at the smaller of the upper thresholds; when both upper thresholds are small, it is again optimal not to use a mixture, and only stop at the larger of the upper thresholds. When the smaller upper threshold is in the range 1 – 3, it is optimal to use a mixed strategy, with most of the mixture distribution on the smaller of the two upper thresholds. Again, $X_0 = x = 0.2$. The optimal mixture is to place probability mass 0.75 on threshold 1.1 and weight 0.25 on threshold 3.1. The family of S -shaped utility functions u_i as defined in (3.20) are used. Parameters used are $\psi = 1/2$ for the price process, $N = 3$ and $\{(\delta_i, R_i, \kappa_i)\} = \{(0.15, 1, 2), (0.1, 2, 2), (0.08, 3, 2)\}$ for the utility functions where for each i , $\delta_i \in (0, 1)$ is a risk aversion/risk seeking parameter, $R_i > 0$ is the reference level and $\kappa_i \geq 1$ is the loss aversion parameter.

diminishing, and the improvement in the CSC value from allowing mixed strategies which randomise over 4 upper thresholds is negligible.

The results of randomisation among upper thresholds for the family of S -shaped utility functions (in Figure 3.3) are summarized in Table 3.1.

3.6.2 An example based on concave utilities

In the previous example we used a family of S -shaped reference dependent utility functions. We saw that the problem could be reduced to a problem which shared many features with the stylized problem of Section 3.5. In this section we build a model using concave utility functions. We build our example from the sum of a

Number of Thresholds	Best Thresholds	Best Mass Distribution	Best CSC
1	2.75	1	0.7263
2	(1.1, 3.1)	(0.75, 0.25)	0.8368
3	(1.1, 2.1, 3.1)	(0.76, 0.11, 0.13)	0.8425
4	Negligible improvement over 3 thresholds case		

Table 3.1: Summary of results of randomisation among upper thresholds for the family of S-shaped utility functions in Figure 3.3.

Number of Thresholds	Best Thresholds	Best Mass Distribution	Best CSC
1	22..68	1	0.6215
2	(3.84, 187.42)	(0.56, 0.44)	0.6373
3	Negligible improvement over 2 thresholds case		

Table 3.2: Summary of results of randomisation among upper thresholds for the family of concave utility functions in Figure 3.8.

power utility function and an exponential utility.²

As in Section 3.6.1, suppose Y is geometric Brownian motion with scale function $s(z) = z^\psi$ for $\psi \in (0, 1)$. For γ, κ, ϕ non-negative constants, define $f = f_{\gamma, \kappa, \phi} : \mathbb{R}^+ \mapsto \mathbb{R}^+$ by

$$f(z) = z^\psi + \frac{1}{\gamma} - \frac{1}{\gamma} e^{-\gamma[(z+\kappa)^\phi - \kappa^\phi]} \quad (3.25)$$

Then $f(0) = 0$ and provided $\phi < 1$, f is concave. Set $g(w) = f \circ s^{-1}$ so that $g_{\gamma, \kappa, \phi}(w) = f_{\gamma, \kappa, \phi}(w^{1/\psi})$. Then

$$g(w) = w + \frac{1}{\gamma} - \frac{1}{\gamma} e^{-\gamma[(w^{1/\psi} + \kappa)^\phi - \kappa^\phi]}. \quad (3.26)$$

Provided $\psi < \phi$ we have that g is convex for small values of w and concave for larger values. We will thus assume $\psi < \phi < 1$.

Let $\mathcal{W}^Y = \{u_i : 1 \leq i \leq N\}$ where $u_i(z) = f_{\gamma_i, \kappa_i, \phi_i}(z)$. Then $\mathcal{W}^X = \{g_i : 1 \leq i \leq N\}$ where $g_i(w) = g_{\gamma_i, \kappa_i, \phi_i}(w)$. Families of functions \mathcal{W}^Y and \mathcal{W}^X are given in Figure 3.8 for the parameters:

$$\psi = 1/4, N = 3 \text{ and } \{(\gamma_i, \kappa_i, \phi_i)\} = \{(0.9, 1, 0.9), (0.5, 10, 0.4), (0.2, 20, 0.3)\}. \quad (3.27)$$

²Our experience is that it is quite difficult to build examples based on families of concave utilities for which randomisation is beneficial, especially if we restrict attention to standard one-parameter families (eg. CRRA or CARA). However, this example shows that it is possible to build examples of families of concave utilities for which randomisation over thresholds is beneficial.

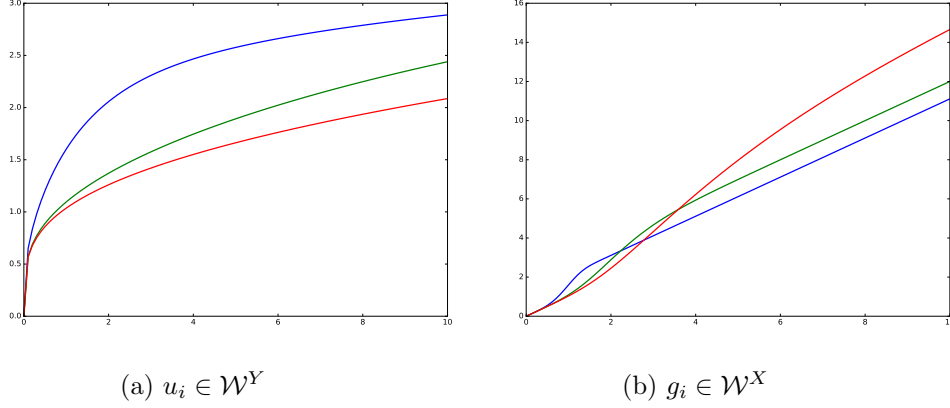


Figure 3.8: The families of concave utility functions $\mathcal{W}^Y = \{u_i : 1 \leq i \leq N\}$ with $u_i(z) = f_{\gamma_i, \kappa_i, \phi_i}(z)$ and f defined in (3.25). In natural scale, $\mathcal{W}^X = \{g_i : 1 \leq i \leq N\}$ with $g_i(w) = g_{\gamma_i, \kappa_i, \phi_i}(w)$ where g is given in (3.26). Parameters used are $\psi = 1/4$ for the price process, $N = 3$ and $\{(\gamma_i, \kappa_i, \phi_i)\} = \{(0.9, 1, 0.9), (0.5, 10, 0.4), (0.2, 20, 0.3)\}$.

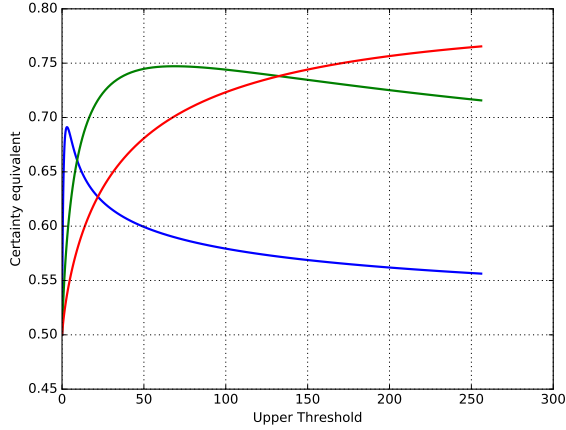


Figure 3.9: The certainty equivalent value under a pure threshold strategy $\tau_{0,\gamma}^X = \inf\{t : X_t \notin (0, \gamma)\}$ as a function of upper threshold γ for $\gamma > X_0 = x = 0.5$. The family of concave utility functions $u_i(z) = f_{\gamma_i, \kappa_i, \phi_i}(z)$ with f defined in (3.25) are used. The best pure threshold strategy uses an upper threshold of approximately 22.68 and gives a CSC certainty equivalent of 0.6215, as marked on the figure. Parameters used are $\psi = 1/4$ for the price process, $N = 3$ and $\{(\gamma_i, \kappa_i, \phi_i)\} = \{(0.9, 1, 0.9), (0.5, 10, 0.4), (0.2, 20, 0.3)\}$.

Consider first pure threshold strategies, $\tau_{0,\gamma}^X$ for different upper thresholds γ . The certainty equivalents associated with the utilities $\{u_i\}_{i=1,2,3}$ as a function of the upper threshold are plotted in Figure 3.9 with an initial value of $X_0 = x = 0.5$. We

see from the figure that the best pure threshold strategy uses an upper threshold of approximately 22.68 and yields a CSC certainty equivalent of 0.6215.

If we now search for the best randomisation over two upper thresholds we find that the best strategy is to assign probability mass 0.56, 0.44 to thresholds 3.84, 187.42 respectively and that this gives a CSC value of 0.6373. Again the best randomised strategy is strictly better than any pure threshold strategy. However, allowing randomisation over three upper thresholds brings only negligible further benefits. The results of randomisation among upper thresholds for the family of concave utilities (in Figure 3.8) are summarized in Table 3.2.

3.7 Conclusion

This paper considers agents who exhibit cautious stochastic choice (CSC) and who face optimal timing or stopping decisions in a dynamic setting. We build on the seminal work on CSC in a static setting by Cerreia-Vioglio et al (2015, 2017) and provide a continuous-time optimal stopping model under CSC. In our dynamic setup, the value associated with a stopping rule is not quasi-convex and hence we cannot necessarily expect there to be a pure threshold rule which is optimal. Despite this observation, it is quite a challenge to find examples where it can be clearly demonstrated that the optimal stopping rule is a non-trivial mixture of threshold strategies. This paper has taken up this challenge and provides first, a stylized, tractable example whereby the optimal stopping rule and value can be constructed explicitly, and second, two realistic example models under reference-dependent or concave families of utility functions under which pure threshold strategies are not optimal. Our predictions are in line with recent experimental evidence in dynamic settings whereby individuals do not play cut-off or threshold strategies (Strack and Viefers (2017), Fischbacher, Hoffmann and Schudy (2015)).

Chapter 4

Optimal Stopping and the Sufficiency of Randomised Threshold Strategies

4.1 Introduction and main results

Let $Y = (Y_t)_{t \geq 0}$ be a one-dimensional, time-homogeneous, continuous strong-Markov process. Let \mathcal{T} be the set of all stopping times, let \mathcal{T}_T be the set of all threshold stopping times, and let \mathcal{T}_R be the set of randomised threshold stopping times (see definition in Chapter 3.4.1). Note that $\mathcal{T}_T \subset \mathcal{T}_R \subset \mathcal{T}$. Let $V = V(\tau)$ be the value associated with a stopping rule τ . Consider the optimal stopping problem associated with V , ie. the problem of finding

$$V_*(\mathcal{S}) = \sup_{\tau \in \mathcal{S}} V(\tau) \quad (4.1)$$

where \mathcal{S} is some set of stopping times (for example $\mathcal{S} = \mathcal{T}$ or $\mathcal{S} = \mathcal{T}_T$), and especially the problem of finding an optimizer for (4.1). Let $Q(\mathcal{S}) = \{\mu : \mu = \mathcal{L}(Y_\tau), \tau \in \mathcal{S}\}$. where $\mathcal{L}(Y_\tau)$ is the law of Y_τ .

Assumption 2 (Law invariance). *V is law invariant, ie $V(\tau) = H(\mathcal{L}(Y_\tau))$ for some function $H : Q(\mathcal{T}) \mapsto \mathbb{R}$.*

We say that $V = V(\tau)$ is law invariant if, whenever σ, τ are stopping times, $\mathcal{L}(Y_\sigma) = \mathcal{L}(Y_\tau)$ implies that $V(\sigma) = V(\tau)$.

Recall that H is quasi-convex if $H(\lambda\mu_1 + (1 - \lambda)\mu_2) \leq \max\{H(\mu_1), H(\mu_2)\}$ for $\mu_{1,2} \in Q(\mathcal{T})$ and $\lambda \in (0, 1)$. It follows by induction that if $\mu = \sum_{i=1}^N \lambda_i \mu_i$ where

$\lambda_i \geq 0$, $\sum_{i=1}^N \lambda_i = 1$ and $\mu_i \in Q(\mathcal{T})$ then

$$H(\mu) \leq \max_{1 \leq i \leq N} H(\mu_i) \leq \sup_{\tilde{\mu} \in Q(\mathcal{T})} H(\tilde{\mu}). \quad (4.2)$$

Recall also that if H is lower semi-continuous and $\mu_n \Rightarrow \mu$ then $H(\mu) \leq \liminf H(\mu_n)$. It follows that $H(\mu) \leq \limsup H(\mu_n)$.

The following result is well-known, but we include it as a contrast to our result on the sufficiency of randomised threshold rules.

Main Result 1 (See Theorem 8 below). *Suppose H is quasi-convex and lower semi-continuous. Then $V_*(\mathcal{T}_T) = V_*(\mathcal{T})$.*

Corollary 3. *In the setting of Theorem 8, in solving the optimal stopping problem (4.1) over the set of all stopping times it is sufficient to restrict attention to threshold rules.*

As the canonical example, consider expected utility where the utility is represented by a continuous, increasing function u . Then, $V(\tau) = \mathbb{E}[u(Y_\tau)]$, assuming that the expectation is well defined. It follows that V is law invariant. Indeed $V(\tau) = H(\mathcal{L}(Y_\tau))$ where $H(\zeta) = \int u(z)\zeta(dz)$. H is quasi-convex and lower semi-continuous. In this example it is well known that there is an optimal stopping rule which is of threshold form, see for example, Dayanik and Karatzas (2003). The fact that quasi-convexity means that there is no benefit from following randomised strategies is well understood in the economics literature, see Machina (1985) Camerer and Ho (1994), Wakker (2010) and He et al (2017).

Recently there has been a surge of interest in problems which, whilst they have the law invariance property, do not satisfy the quasi-convex criterion. Two examples are optimal stopping under prospect theory (Xu and Zhou (2013)), and optimal stopping under cautious stochastic choice (Henderson et al (2017) [23]).

Introduce the set \mathcal{T}_R of mixed or randomised threshold rules (i.e. stopping rules which are based on the first exit from a randomly chosen interval).

Main Result 2 (See Theorem 7 below). *Suppose law invariance holds for V , but not quasi-convexity for H . Then $V_*(\mathcal{T}_T) \leq V_*(\mathcal{T}_R) = V_*(\mathcal{T})$.*

We will show by example that the first inequality may be strict.

Corollary 4. *In the setting of Theorem 7, in solving the optimal stopping problem (4.1) over the set of all stopping rules it is sufficient to restrict attention to randomised threshold rules, but it may not be sufficient to restrict attention to (pure) threshold rules.*

It should be noted that we do not include discounting in our analysis since a problem involving discounting does not satisfy the law invariance property. Nonetheless, as is well known, the conclusion of Corollary 3 remains true for the problem of maximizing discounted expected utility of the stopped process $V(\tau) = \mathbb{E}[e^{-\beta\tau}u(Y_\tau)]$. However, in problems which go beyond the expected utility paradigm, there are often modelling issues which mitigate against the inclusion of discounting. For this reason, historically the literature has concentrated on problems with no discounting. Finding the optimal stopping rule is often already challenging in these models.

The significance of Corollary 4 is as follows. In many classical models optimal stopping behaviour involves stopping on first exit from an interval (threshold strategies). However, our result implies that the converse is not true: if decision makers are observed to stop only when the process is reaching new maxima or minima, then it does not necessarily mean that they are using threshold strategies. Instead the decision criteria may be more complicated, and they may be utilising a randomised threshold rule.

4.2 Problem specification and the problem in natural scale

We work on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$. Let $Y = (Y_t)_{t \geq 0}$ be a (\mathbb{F}, \mathbb{P}) -stochastic process on this probability space with state space I which is an interval. Let \bar{I} be the closure of I . We suppose that Y is a regular, one-dimensional, time-homogeneous diffusion with initial value $Y_0 = y$ such that y lies in the interior of I . For Γ an interval in \mathbb{R} or a rectangular set in \mathbb{R}^2 , let $\mathcal{B}(\Gamma)$ denote the Borel σ -algebra on Γ , and let $\mathcal{P}(\Gamma)$ denote the set of probability measures on $(\Gamma, \mathcal{B}(\Gamma))$.

Given that the value associated with a stopping rule is law invariant, one natural approach to finding the optimal stopping time is to try to characterize $Q(\mathcal{S})$. Often, the best way to do this is via a change of scale. Let s be a strictly increasing function such that $X = s(Y)$ is a local martingale. Let $I^X = s(I)$ and let \bar{I}^X be the closure of I^X . Then X is a regular, time-homogeneous local-martingale diffusion on I^X with initial value $x = s(y)$.

Set $Q^X(\mathcal{S}) = \{\nu : \nu = \mathcal{L}(X_\tau), \tau \in \mathcal{S}\}$. Then if $\mathcal{L}(X_\tau) = \nu$ we have $\mathcal{L}(Y_\tau) = \nu \# s$ where $(\nu \# s)(D) = \nu(s(D))$. It follows that $\nu \in Q^X(\mathcal{S})$ if and only if $\nu \# s \in Q(\mathcal{S})$ and hence

$$Q(\mathcal{S}) = \{\nu \# s; \nu \in Q^X(\mathcal{S})\}. \quad (4.3)$$

Thus, if we can characterize $Q^X(\mathcal{S})$ then we can also characterize $Q(\mathcal{S})$. Moreover,

defining $H^X : Q^X(\mathcal{T}) \mapsto \mathbb{R}$ by $H^X(\nu) = H(\nu \sharp s)$ we have $V_*(\mathcal{S}) = \sup_{\mu \in Q(\mathcal{S})} H(\mu) = \sup_{\nu \in Q^X(\mathcal{S})} H^X(\nu)$. The problem of optimizing over stopping laws for the problem with Y becomes a problem of optimizing over the possible laws of the stopped process X in natural scale.

Note that $\tau_{a,b} = \inf_{u \geq 0} \{u : Y_u \notin (a,b)\} = \inf_{u \geq 0} \{u : X_u \notin (s(a), s(b))\} =: \tau_{s(a),s(b)}^X$. Hence \mathcal{T}_T has the alternative representation

$$\mathcal{T}_T = \mathcal{T} \cap \left(\bigcup_{\beta \leq x \leq \gamma; \beta, \gamma \in \bar{I}^X} \{\tau_{\beta, \gamma}^X\} \right),$$

and the set of threshold stopping times for Y is the set of threshold stopping times for X . Similarly, \mathcal{T}_R can be rewritten as $\mathcal{T}_R = \mathcal{T} \cap (\{\tau_\eta^X : \eta \in \mathcal{P}(\mathcal{D}^X)\})$ where $\mathcal{D}^X = ([-\infty, x] \cap \bar{I}^X) \times ([x, \infty) \cap \bar{I}^X)$ and

$$\tau_\eta^X = \inf_{u \geq 0} \{u : X_u \notin (A_\eta, B_\eta) \text{ where } (A_\eta, B_\eta) \text{ has law } \eta\}.$$

4.3 Characterizing the possible laws of the stopped process in natural scale

If $X = s(Y)$ is in natural scale then the state space of X is an interval $I^X = s(I)$ and $X_0 = x := s(y)$. There are four cases:

1. I^X is bounded;
2. I^X is unbounded above but bounded below;
3. I^X is bounded above but unbounded below;
4. I^X is unbounded above and below.

The third case can be reduced to the second by reflection. The first case is generally similar to the second case, and typically the proofs are similar but simpler. The final case is degenerate and will be treated separately. In the main text we will mainly present arguments for the second case (with the other cases covered in Chapter 4.6), but results will be stated in a form which applies in all cases.

Henceforth, in the main text we suppose I^X is bounded below, but unbounded above. Without loss of generality we may assume $I^X = (0, \infty)$ or $[0, \infty)$. Then X is a non-negative local martingale and hence a super-martingale. Moreover, $\lim_{t \rightarrow \infty} X_t$ exists. Hence \mathcal{T} includes stopping rules which take infinite values and on $\{\tau = \infty\}$ we assume $X_\tau = \lim_{t \rightarrow \infty} X_t = 0$. In this case \mathcal{T} is the set of all stopping times

and the intersection with \mathcal{T} in the definitions (3.3) and (3.4) is not necessary. By Fatou's lemma and the super-martingale property

$$\mathbb{E}[X_\tau] = \mathbb{E}[\lim_{t \rightarrow \infty} X_{t \wedge \tau}] \leq \liminf_{t \rightarrow \infty} \mathbb{E}[X_{t \wedge \tau}] \leq x.$$

In particular, if we set $\mathcal{P}_{\leq x} = \{\nu \in \mathcal{P}([0, \infty)) : \int z\nu(dz) \leq x\}$ then $Q^X(\mathcal{T}) \subseteq \mathcal{P}_{\leq x}$.

Lemma 7. $Q^X(\mathcal{T}) = Q^X(\mathcal{T}_R)$.

Proof. Here we prove the lemma in the case where I^X is bounded below. We show that $Q^X(\mathcal{T}) = Q^X(\mathcal{T}_R) = \mathcal{P}_{\leq x}$. Given $\nu \in \mathcal{P}_{\leq x}$ the aim is to find a stopping time $\tau \in \mathcal{T}_R$ such that $\mathcal{L}(X_\tau) = \nu$. The task of finding general stopping times with $\mathcal{L}(X_\tau) = \xi$ for given $\xi \in \mathcal{P}(\bar{I}^X)$ is known as the Skorokhod embedding problem (Skorokhod (1965)). In fact we use an extension of an embedding due to Hall (1998), see also Durrett (1991). The extension relates to the fact that we allow for target laws which have a different mean to the initial value of X , whereas the Hall embedding assumes $\int z\nu(dz) = x$. The Hall embedding, and the extension we give, are mixtures of threshold strategies.

Suppose ν is an element of $\mathcal{P}_{\leq x}$ (and ν is not a point mass at x). The case of $\nu = \delta_x$ corresponds to the (threshold) stopping time $\tau = 0$. Let G be the (right-continuous) quantile function of ν . We have $x \geq \int z\nu(dz) = \int_{(0,1)} G(u)du$. In particular, unless $\lim_{u \uparrow 1} G(u) \leq x$ there exists a unique solution $v^* \in [0, 1)$ to $\int_v^1 [G(w) - x]dw = 0$. Let $z^* = G(v^*) \leq x$. If $\lim_{u \uparrow 1} G(u) \leq x$ then set $v^* = 1$ and $z^* = \lim_{u \uparrow 1} G(u)$.

Let ν_0 be the measure of size v^* such that $\nu_0([0, z]) = v^* \wedge \nu([0, z])$. Then ν_0 has support contained in $[0, z^*]$. Let ν_1 be the measure of size $1 - v^*$ such that $\nu_1([0, z]) = (\nu([0, z]) - v^*)^+$. Then ν_1 has support in $[z^*, \infty)$ and barycentre x . Moreover $\nu = \nu_0 + \nu_1$.

Define $c = \int_x^\infty (y - x)\nu(dy)$. By construction, $c = \int_x^\infty (y - x)\nu_1(dy)$ and we have from the fact that ν_1 has barycentre x that $\int_{z^*}^\infty (y - x)\nu_1(dy) = 0$ and hence

$$c = \int_{z^*}^x (x - y)\nu_1(dy). \quad (4.4)$$

Let $\eta \in \mathcal{P}([0, x] \times (x, \infty])$ be given by

$$\eta(da, db) = \nu_0(da)I_{\{0 \leq a \leq z^*\}}I_{\{b = \infty\}} + \nu_1(da)\nu_1(db)\frac{(b - a)}{c}I_{\{z^* \leq a \leq x < b < \infty\}}.$$

Note first that η is a probability measure:

$$\begin{aligned}
& \int_{0 \leq a \leq x} \int_{x < b \leq \infty} \eta(da, db) \\
&= v^* + \int_{z^* \leq a \leq x} \nu_1(da) \int_{x < b < \infty} \frac{b-x}{c} \nu_1(db) + \int_{z^* \leq a \leq x} \frac{x-a}{c} \nu_1(da) \int_{x < b < \infty} \nu_1(db) \\
&= v^* + \int_{z^* \leq a \leq x} \nu_1(da) + \int_{x < b < \infty} \nu_1(db) = v^* + \nu_1([z^*, \infty)) = 1
\end{aligned}$$

where we use the definition of c and (4.4) in going from the second line to the third.

It remains to show that $\mathcal{L}(X_{\tau_\eta^X}) = \nu$. Let f be a bounded test function. Then, using the fact that if $b = \infty$ then $X_{\tau_{a,\infty}^X} = a$, and the definition of c and (4.4) for the penultimate line,

$$\begin{aligned}
\mathbb{E}[f(X_{\tau_\eta^X})] &= \int \int \eta(da, db) \mathbb{E}[f(X_{\tau_{a,b}^X})] \\
&= \int \nu_0(da) f(a) \\
&\quad + \int_{z^* \leq a \leq x} \int_{x < b < \infty} \nu_1(da) \nu_1(db) \frac{b-a}{c} \left[f(a) \frac{(b-x)}{b-a} + f(b) \frac{(x-a)}{b-a} \right] \\
&= \int \nu_0(da) f(a) + \int_{z^* \leq a \leq x} \nu_1(da) f(a) \int_{x < b < \infty} \nu_1(db) \frac{(b-x)}{c} \\
&\quad + \int_{z^* \leq a \leq x} \frac{(x-a)}{c} \nu_1(da) \int_{x < b < \infty} f(b) \nu_1(db) \\
&= \int_{0 \leq z \leq z^*} f(z) \nu_0(dz) + \int_{z^* \leq z \leq x} f(z) \nu_1(dz) + \int_{x < z} f(z) \nu_1(dz) \\
&= \int f(z) \nu(dz).
\end{aligned}$$

Hence $\mathcal{L}(X_{\tau_\eta^X}) = \nu$ as required. \square

Let $\chi_{a,b} = \frac{b-x}{b-a} \delta_a + \frac{x-a}{b-a} \delta_b$. Then $\chi_{a,b}$ is the law of $X_{\tau_{a,b}^X}$. Moreover, $\mathcal{L}(X_{\tau_{a,\infty}^X}) = \delta_a$. Then,

$$Q^X(\mathcal{T}_T) = (\cup_{0 \leq a \leq x} \delta_x) \cup (\cup_{0 \leq a < x < b < \infty} \chi_{a,b}).$$

4.4 Sufficiency of mixed threshold rules

Our main result is that in a large class of problems it is sufficient to search over the class of mixed threshold rules.

Theorem 7. *Suppose Y is a regular, time-homogeneous diffusion. Suppose the law invariance property holds (Assumption 2) and that the filtration is sufficiently rich.*

Then $V_*(\mathcal{T}) = V_*(\mathcal{T}_R)$.

Proof. Since $Q^X(\mathcal{T}) = Q^X(\mathcal{T}_R)$ (Lemma 7) we have $Q(\mathcal{T}) = Q(\mathcal{T}_R)$. Then

$$V_*(\mathcal{T}) = \sup_{\mu \in Q(\mathcal{T})} H(\mu) = \sup_{\mu \in Q(\mathcal{T}_R)} H(\mu) = V_*(\mathcal{T}_R).$$

□

Note that it is not our claim that every optimal stopping rule is a mixed threshold rule. Typically, at least in the case where $V(\mathcal{T}_T) < V(\mathcal{T})$, there will be other optimal stopping rules which are not of threshold type.

4.4.1 An example - Rank dependent utility

Let Z be a non-negative random variable. Let $v : [0, \infty) \mapsto [0, \infty)$ be an increasing, differentiable function with $v(0) = 0$. Then the expected value of $v(Z)$ can be expressed as $\mathbb{E}[v(Z)] = \int_0^\infty v'(z) (1 - F_Z(z)) dz$, where F_Z is the cumulative distribution function of Z . Under rank-dependent utility (Quiggin [49]) or probability weighting (Tversky and Kahneman [59]) the prospect value $\mathcal{E}_v(Z)$ of Z is

$$\mathcal{E}_v(Z) = \int_0^\infty v'(z) w(1 - F_Z(z)) dz$$

where $w : [0, 1] \mapsto [0, 1]$ is an increasing, differentiable probability weighting function. Writing $G_Z = F_Z^{-1}$ for the quantile function of Z , then after a change of variable and integration by parts we have (see Xu and Zhou [62, Lemma 3.1]) the alternative representation

$$\mathcal{E}_v(Z) = \int_0^1 w'(1 - u) v(G_Z(u)) du.$$

Now let $Y = (Y_t)_{t \geq 0}$ be a non-negative diffusion and consider the problem of maximizing over stopping times the prospect value of the stopped process Y , ie of finding

$$\sup_{\tau \in \mathcal{T}} \mathcal{E}_v(Y_\tau). \tag{4.5}$$

Clearly the prospect value depends on the stopping time only through the law of the stopped process. Hence it is sufficient to characterize the optimal target distribution, for example via its quantile function. Xu and Zhou [62] solve for the optimal quantile function in several cases. One relevant case is the following:

Proposition 5 (Xu and Zhou [62]). *Suppose Y is in natural scale and has state*

space $[0, \infty)$ and initial value y . Suppose v and w are concave. Suppose there exists $\lambda^* \in (0, \infty)$ which solves

$$\int_0^1 (v')^{-1} \left(\frac{\lambda^*}{w'(1-u)} \right) du = y.$$

Then the quantile function of the optimal stopping distribution is given by $G^*(u) = (v')^{-1} \left(\frac{\lambda^*}{w'(1-u)} \right)$.

Proof of Proposition 5. A proof is given in Xu and Zhou [62, Theorem 5.1], but since it is short, elegant and pertinent to our main results we include it here. From the characterization of $Q(\mathcal{T})$ we have that a quantile function must satisfy $\int_0^1 G(u) du \leq y$. By construction G^* has this property, and since v' and w' are decreasing, G^* is increasing. Hence G^* has the properties required of a quantile function of a distribution which can be obtained by stopping Y . On the other hand, for any non-negative function G with $\int_0^1 G(u) du \leq y$,

$$\begin{aligned} \int_0^1 w'(1-u)v(G(u))du &= \int_0^1 [w'(1-u)v(G(u)) - \lambda^* G(u)]du + \lambda^* \int_0^1 G(u)du \\ &\leq \int_0^1 \sup_{g>0} [w'(1-u)v(g) - \lambda^* g]du + \lambda^* y \\ &= \int_0^1 [w'(1-u)v(G^*(u)) - \lambda^* G^*(u)]du + \lambda^* y \\ &= \int_0^1 w'(1-u)v(G^*(u))du. \end{aligned}$$

□

Xu and Zhou (2013) point out that although there is a unique optimal prospect there are infinitely many stopping rules which attain this prospect. They advocate the use of the stopping rule based on the Azéma-Yor (1979) stopping time, in which case the stopping rule has a drawdown feature, and involves stopping the first time the process falls below some function of the maximum. However, by Theorem 7 there is also a randomised threshold rule which is optimal.

4.5 Sufficient conditions for the optimality of pure threshold rules

In this section we argue that if the value associated with a stopping rule is law invariant, and if H is quasi-convex and lower semi-continuous then pure threshold

rules are optimal.

Lemma 8. *Suppose $\nu \in Q^X(\mathcal{T})$ consists of finitely many atoms. Then there exists $\eta \in \mathcal{P}(\mathcal{D}^X)$ such that η consists of finitely many atoms and $\mathcal{L}(X_{\tau_\eta^X}) = \nu$.*

Proof. It follows from the construction in the proof of Lemma 7 that if μ is purely atomic then so is η . \square

Lemma 9. *Let ν be an element of $Q^X(\mathcal{T})$. Then there exist $(\eta_n)_{n \geq 1}$ such that η_n has finite support for each n and such that $\mathcal{L}(X_{\tau_{\eta_n}^X}) \Rightarrow \nu$.*

Proof. Since $\nu \in Q^X(\mathcal{T}) = Q^X(\mathcal{T}_R)$ there exists η such that $\mathcal{L}(X_{\tau_\eta^X}) = \nu$. Let $(\eta_n)_{n \geq 1}$ be a sequence of measures with finite support such that $\eta_n \Rightarrow \eta$. Then for $f : [0, \infty) \mapsto \mathbb{R}$ a bounded continuous test function define $\tilde{f} : [0, x] \times [x, \infty)$ by $\tilde{f}(a, b) = f(a) \frac{b-x}{b-a} + f(b) \frac{x-a}{b-a}$ for $a < b$ with $\tilde{f}(x, x) = f(x)$. Then, since \tilde{f} is bounded and continuous

$$\mathbb{E}[f(X_{\tau_{\eta_n}^X})] = \int \int \eta_n(da, db) \tilde{f}(a, b) \rightarrow \int \int \eta(da, db) \tilde{f}(a, b) = \mathbb{E}[f(X_{\tau_\eta^X})]$$

and it follows that $\nu_n := \mathcal{L}(X_{\tau_{\eta_n}^X}) \Rightarrow \nu$. \square

Theorem 8. *Suppose Y is a regular, time-homogeneous diffusion. Suppose the law invariance property holds (Assumption 2). Suppose that H is quasi-convex and lower semi-continuous. Then $V_*(\mathcal{T}) = V_*(\mathcal{T}_T)$.*

Proof. Clearly $V_*(\mathcal{T}) \geq V_*(\mathcal{T}_T)$.

For any μ_n with finite support we can define $\nu_n = \mu_n \# s^{-1}$. Then we can find a measure η_n with finite support such that $\mathcal{L}(X_{\tau_{\eta_n}^X}) = \nu_n$. Moreover ν_n can be decomposed as a convex combination

$$\nu_n = \sum_{i=1}^N \gamma_i \chi_{a_i, b_i} + \sum_{j=1}^M \lambda_j \delta_{a_j}.$$

Then, since H is quasi-convex,

$$\begin{aligned} H(\mu_n) &\leq \left(\max_{1 \leq i \leq N} H(\chi_{a_i, b_i} \# s^{-1}) \right) \vee \left(\max_{1 \leq j \leq M} H(\delta_{s^{-1}(a_j)}) \right) \\ &\leq \left(\sup_{0 \leq a \leq x \leq b < \infty} H(\chi_{a, b} \# s^{-1}) \right) \vee \left(\sup_{0 \leq a \leq x} H(\delta_{s^{-1}(a)}) \right) = V_*(\mathcal{T}_T). \end{aligned}$$

Then, for $\tau \in \mathcal{T}$, if $\mu = \mathcal{L}(Y_\tau)$ and if $\mu_n \Rightarrow \mu$

$$V(\tau) = H(\mu) \leq \limsup H(\mu_n) \leq V_*(\mathcal{T}_T).$$

Hence $V_*(\mathcal{T}) \leq V_*(\mathcal{T}_T)$. □

4.6 Extension to other state spaces for the process in natural scale

4.6.1 The range is unbounded below but bounded above

In this case we may assume that $I^X = (-\infty, 0)$ or $(-\infty, 0]$. The analysis goes through almost unchanged except that now X is a convergent sub-martingale and $Q^X(\mathcal{T}) = Q(\mathcal{T}_R) = \mathcal{P}_{\geq x}$ where $\mathcal{P}_{\geq x} = \{\nu \in \mathcal{P}((-\infty, 0]) : \int z\nu(dz) \geq x\}$.

4.6.2 The range is bounded

Suppose X is bounded. In this case $Q(\mathcal{T}) = Q(\mathcal{T}_R) = \mathcal{P}_{=x}$ where $\mathcal{P}_{=x} = \{\nu \in \mathcal{P}(\bar{I}^X) : \int z\nu(dz) = x\}$. To see this note that X is a uniformly integrable martingale and not just a super-martingale. Therefore we must have $\mathbb{E}[X_\tau] = \lim \mathbb{E}[X_{\tau \wedge t}] = x$ and hence $Q(\mathcal{T}) \subseteq \mathcal{P}_{=x}$. Conversely, by the same argument as in Lemma 7, but this time with $v^* = 0$ and $\nu_1 \equiv \nu$, we deduce that for any $\nu \in \mathcal{P}_{=x}$ there exists a randomisation η such that $\mathcal{L}(X_{\tau_\eta^X}) = \nu$. It follows that $Q(\mathcal{T}) = Q(\mathcal{T}_R) = \mathcal{P}_{=x}$.

The proofs of Lemma 8, Lemma 9 and Theorem 7 go through unchanged.

4.6.3 The range is unbounded above and below

Now suppose I^X is unbounded above and below. By the Rogozin trichotomy (Rogozin (1996)) $-\infty = \liminf_t X_t < x < \limsup_t X_t = \infty$ and $\lim_{t \uparrow \infty} X_t$ does not exist. In this case we must restrict \mathcal{T} to the set of stopping times with $\mathbb{P}(\tau < \infty) = 1$. In the main text we set $\mathcal{T}_T = \mathcal{T} \cap (\cup_{\beta \leq y \leq \gamma, \beta, \gamma \in \bar{I}^Y} \{\tau_{\beta, \gamma}\})$ but we could equivalently write $\mathcal{T}_T = \cup_{(\beta, \gamma) \in \mathcal{D}_0} \{\tau_{\beta, \gamma}\}$, where $\mathcal{D}_0 = ([-\infty, y] \cap \bar{I}^Y) \times ([y, \infty] \cap \bar{I}^Y) \setminus \{s^{-1}(-\infty), s^{-1}(\infty)\}$. We have to exclude the threshold rule $\tau_{s^{-1}(-\infty), s^{-1}(\infty)}$ since $\tau_{s^{-1}(-\infty), s^{-1}(\infty)} = \infty$ almost surely and Y_∞ is not defined. In terms of threshold rules $\tau_{a,b}^X$ for X we allow $a = -\infty$ or $b = \infty$ but not both. Then $\mathcal{T}_T = \{\tau_{\beta, \gamma} : (\beta, \gamma) \in \mathcal{D}_0^X\}$ where $\mathcal{D}_0^X = \mathcal{D}^X \setminus \{-\infty, \infty\} = [\infty, x] \times [x, \infty] \setminus \{-\infty, \infty\}$.

In the definition of randomised threshold rules we can write $\mathcal{T}_R = \{\tau_\zeta : \zeta \in \mathcal{P}(\mathcal{D}_0)\}$ where \mathcal{D}_0 is as above and similarly $\mathcal{T}_R = \{\tau_\eta^X : \eta \in \mathcal{P}(\mathcal{D}_0^X)\}$.

When $I^X = \mathbb{R}$ we claim that we have $Q^X(\mathcal{T}) = Q^X(\mathcal{T}_R) = \mathcal{P}(\mathbb{R})$. Since stopping times are finite almost surely we must have $Q^X(\mathcal{T}) \subseteq \mathcal{P}(\mathbb{R})$ so it is sufficient to show that for any $\nu \in \mathcal{P}(\mathbb{R})$ we have $\nu \in Q^X(\mathcal{T}_R)$. Given $\nu \in \mathcal{P}(\mathbb{R})$ let A_ν be a \mathcal{F}_0 -measurable random variable with law ν and set $\tau = \inf\{u : X_u = A_\nu\}$. Then $\mathcal{L}(X_\tau) = \mathcal{L}(A_\nu) = \nu$.

The proofs of Lemma 8, Lemma 9 and Theorem 7 go through unchanged.

4.7 Conclusion

In classical optimal stopping problems involving maximizing expected utility the optimal strategy is a threshold rule and involves stopping the first time that the process leaves an interval. However, in more general settings the optimal strategy may be more sophisticated. In some settings, for example those involving regret (Loomes and Sugden (1982)) the optimal stopping rule may depend on some functional of the path (for example the maximum price to date). But, as argued here, for a large class of problems the payoff depends only on the distribution of the stopped process, and then there are many optimal stopping rules, some of which take the form of randomized threshold rules. In this article we have utilized (an extended version of) the Hall solution of the Skorokhod embedding problem (Hall (1985)) to give our randomized threshold rule, but there are other solutions of the Skorokhod embedding which can also be viewed as mixed threshold rules, including the original solution of Skorokhod (1965) and the solution of Hirsch et al (2011).

The idea that if the objective is expressed in terms of a function which is not quasi-convex then agents may want to use randomized strategies is well appreciated in static settings. In a dynamic setting He et al (2017) argue that in binomial-tree, probability-weighted model of a casino (Barberis (2014)) gamblers may prefer path-dependent strategies over strategies which are defined via a partition of the set of nodes into those at which the gambler stops and those at which he continues. (See also Ebert and Strack (2016) and Henderson et al (2017) for discussion of a related optimal stopping problem with probability weighting based on a diffusion process.) He et al (2017) argue further that the path-dependent strategy can be replaced by a randomized strategy under which the decision about whether to stop at a node depends not on the path history but rather the realization of an independent uniform random variable. This preference for randomization mirrors our result, but takes a different form. In our perpetual problem the agent chooses a randomized pair of levels and then follows a threshold strategy based on these levels. In He et al (2017) a zero-one decision about whether to stop at a node is replaced by a probability of continuing, and the stopping rules which arise are not randomized threshold rules.

Many optimal stopping models in the economics literature predict that the agent will stop on first exit from an interval, which necessarily involves stopping either at the current maximum or the current minimum. If instead, observed behavior includes stopping at levels which are not equal to one of the running extrema of the

process then this is evidence against the model (Strack and Viefers (2017) present experimental evidence from a laboratory game that players do not follow threshold strategies - instead players visit the same price three times on average before stopping). But, our results imply that the converse is not true. Even if agents only ever take a decision to sell at a time when the process is at a new maximum or new minimum, this does not necessarily mean that agents are following a pure threshold rule. They could have any target distribution, as for example in Proposition 5, but be realizing this target distribution via a randomized threshold rule.

Chapter 5

Randomising Rules for Stopping Problems

5.1 Introduction

In a classical, continuous-time, optimal stopping problem the agent chooses the best time to stop a stochastic process in order to maximise the expected discounted return. The agent can choose when to stop and if at any moment they decide to stop, stopping occurs immediately with probability one. However, in many settings this is an idealistic oversimplification. There may be many reasons why agents do not make an unequivocal best choice when choosing between stopping and continuing. For example, they may be unable to precisely evaluate the value of continuing (or alternatively have imprecise information about the value of stopping), they may be unable to put their stopping decision into practice (they may wish to sell, but find no buyer) or they may have an ulterior motive for not choosing the apparently best option (perhaps they delay sale to learn more about alternative outcomes). Following Strack and Viefers (2017) we consider a modification of the problem in which stopping occurs at a rate which depends on the relative values of stopping and continuing (i.e. stopping probability at any decision point is given by $p = \frac{g}{g+c}$ with g being the stopping value and c being the perceived continuation value).

One immediate issue is that if at each instant an agent has a positive probability of stopping, then since in a continuous-time model there are an uncountable number of stopping opportunities in any small interval, the agent will end up stopping immediately. To deal with the first issue, we constrain the agent to stopping at a countable number of times, namely the event times $\{T_n^\lambda\}_{n \geq 1}$ of an exogenous Poisson process. For our purposes the memoryless property of the Poisson process is crucial

in allowing us to conclude that the value function is a Markovian function of the state process, which keeps the analysis tractable.

Another issue we need to address is how to define the continuation value. Our inspiration is a paper by Strack and Viefers (2017) who analyse a stopping decision under a randomising stopping rule. They take the perceived continuation value as the value under optimal stopping rule. This situation models an agent who can determine the optimal stopping rule, but cannot ensure that the optimal rule is followed exactly; such agent is not sophisticated enough to allow for the fact that their future self will not behave optimally. The innovation of this thesis is that we introduce a new type of randomised stopping in which the perceived continuation value is calculated based on the fact that stopping will be determined by the randomised rule. This models an agent who is aware that their future self is not able to stop optimally, but rather stops with a randomised rule, and who values the problem accordingly. This definition introduces non-linear feedback into the valuation problem.

Nevertheless, we solve the randomised stopping problem for different specifications of the perceived continuation value. We will also give various alternative characterisations of the solution including a stochastic representation and a representation as the solution of linear growth of an ordinary differential equation.

Our final set of findings concern the case in which the rate of the Poisson process describing opportunities to stop increases to infinity. We show that it is possible to choose the stopping probability in such a way that the problem has a non-degenerate limit. Then we give a description of a continuous-time stopping problem for which the value function solves the identical equation to the aforementioned limiting problem. This newly introduced problem involves stopping at the first event time of an inhomogeneous stopping time with rate depending on the ratio of the instantaneous stopping value to the continuation value.

5.2 Problem Specification

Recall the stopping problems described in Chapter 1. Let the asset price process $X = (X_t)_{t \geq 0}$ be a time-homogeneous, continuous, real-valued, strong-Markov process with initial value $X_0 = x$, living on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = \{\mathcal{F}_t\})$ which satisfies the usual conditions. Let $g : \mathbb{R} \mapsto \mathbb{R}_+$ be a (measurable) payoff function (satisfying suitable growth conditions, so that the problem is well-posed) and let β be a strictly positive discount factor. The value function $w = w(x)$ of the

classical discounted optimal stopping problem is defined by (1.1)

$$w(x) = \sup_{\tau \in \mathcal{T}} \mathbb{E}^x[e^{-\beta\tau} g(X_\tau)]$$

where \mathcal{T} is the set of all \mathbb{F} -stopping times.

For a constrained optimal stopping problem in which stopping can only occur at the event times $\{T_n^\lambda\}_{n \geq 1}$ of an independent Poisson process of rate λ , we assume that the probability space is rich enough to carry a Poisson Process which is independent of X , and to carry any other random variables which we wish to define. The value function $h = h(x)$ is now given by (1.5)

$$h(x) = \sup_{\tau \in \mathcal{T}^\lambda} \mathbb{E}^x[e^{-\beta\tau} g(X_\tau)]$$

where \mathcal{T}^λ is the set of all stopping times taking values in the event times of the Poisson process.

Now consider the stopping problem under a randomised stopping rule. We assume that stopping can only occur at event times of a Poisson process. At such an event time the probability of stopping depends on the value of immediate stopping $g = g(X_t)$ and on the perceived continuation value $c = c(X_t)$. More precisely, we suppose there is a map $\Gamma : \mathbb{R}_+ \times \mathbb{R}_+ \mapsto [0, 1]$ such that the probability of stopping is $p(X_t) = \Gamma(g(X_t), c(X_t))$. To simplify notation, we denote $\Gamma^{g,c} = \Gamma(g, c)$.

We can formalise the stopping rule as follows. Suppose the filtration is also sufficiently rich to include $(U_n)_{n \geq 1}$ be a sequence of iid standard uniform random variables with $U_i \in [0, 1]$, which are also independent of X and the Poisson process. Then, at the n^{th} event time of the Poisson process, the conditional probability of stopping is $\mathbb{P}(U_n \leq \Gamma^{g,c}(X_{T_n^\lambda})) = \Gamma^{g,c}(X_{T_n^\lambda})$. Define $\tau_c = T_N^\lambda$ where $N = \min\{n : U_n \leq \Gamma^{g,c}(X_{T_n^\lambda})\}$. Then, the continuation value of the randomised stopping problem is

Problem 1 (Stopping Time Formulation (STF)).

$$G(x) = \mathbb{E}^x[e^{-\beta\tau_c} g(X_{\tau_c})]. \quad (5.1)$$

Integrating against the time of the first event of the Poisson process, and in analogy to (1.7), we have a second formulation for G in feedback form

Problem 2 (Stochastic Formulation with Feedback (SFF)). *G is of linear growth*

and solves

$$G(x) = \mathbb{E}^x \left[\int_0^\infty dt \lambda e^{-\lambda t} e^{-\beta t} \{ \Gamma^{g,c}(X_t) g(X_t) + (1 - \Gamma^{g,c}(X_t)) G(X_t) \} \right]. \quad (5.2)$$

assuming the payoff from immediate stopping is g , the value of continuing is G , and the probability of continuing is $\Gamma^{g,c}$ where c is the perceived value of continuing.

Another way to characterise G is via differential equation. The corresponding representation for G is

Problem 3 (Ordinary Differential Equation Formulation (ODEF)). *G is of linear growth and solves*

$$\mathcal{L}G(x) - \beta G(x) + \lambda \Gamma^{g,c}(x) [g(x) - G(x)] = 0, \quad G(0) = 0. \quad (5.3)$$

The equivalence between (5.1), (5.2) and (5.3) are provided by Lemma 10, Lemma 11 and Lemma 12.

There are several possible choices for the perceived continuation value c . We may take the value of the classical optimal stopping problem w as in Strack and Viefers (2017). Or, given that stopping is only allowed at event times of the Poisson process we can take $c = h$. The novelty in this paper is that we consider the case of a sophisticated agent whose probability of stopping depends on the true continuation value and who takes $c = G$. In this case, since G appears on both sides of (5.1), we need to prove that there exists a fixed point, which is provided in Proposition 6.

Then, in Sections 5.4 and 5.5, we consider solutions to the problem for particular choices of payoff function. First we consider the case $g(x) = x$ when analytic solutions are available. Then we present numerical solutions to the problem when $g(x) = (x - K)^+$ with $K = 1$.

In Section 5.6 we consider what happens in the limit as λ gets large. We show how we can obtain a sensible limit by carefully choosing Γ . Also, we obtain a specification for a continuous time, randomised stopping problem which is non-degenerate.

5.3 The base case

Within the general set-up described above we will mainly work with the following specification.

For the Markov process X we take exponential Brownian motion started at x :

$$\frac{dX_t}{X_t} = \mu dt + \sigma dW_t; \quad X_0 = x.$$

Then X has generator $\mathcal{L} = \mathcal{L}^X$ given by $\mathcal{L}f = \frac{1}{2}x^2\sigma^2f'' + \mu xf'$.

We assume the payoff function g is continuous, non-negative, has at most linear growth, and satisfies $g(0) = 0$. For well-posedness of the classical optimal stopping problem we need $\beta > \mu$ and we assume this parameter restriction throughout.

For the probability of stopping map Γ we take $\Gamma^{g,c} = \frac{g}{g+c}$, although later we will also consider $\Gamma_\xi^{g,c} = \frac{g}{g+\xi c}$ for some weighting parameter ξ .

As a motivation for this choice of Γ , and indeed of randomised stopping, suppose the investor is faced with stopping with reward g or continuing with potential reward c . Suppose however, that there is (multiplicative) measurement error in calculating the rewards so that the investor bases his decision on values \tilde{g} and \tilde{c} where $\tilde{g} = gZ^g$, $\tilde{c} = cZ^c$ and $\{Z^g, Z^c\}$ are a pair of independent (of everything) exponential random variables each with unit rate. Suppose the agent makes a rational decision based on the measured values, in the sense that she stops if $\tilde{g} \geq \tilde{c}$. Then, the probability of stopping is $\mathbb{P}(\tilde{g} \geq \tilde{c}) = \mathbb{P}(Z^g > Z^c \frac{c}{g}) = \frac{g}{g+c} = \Gamma^{g,c}$.

The value of the randomised stopping problem is bounded above by the value of the optimal stopping problem (1.1). Since g is of linear growth (and the discount factor is larger than the mean growth rate) w grows at most linearly. Hence also, the solution G is also of linear growth.

In this section we concentrate on the extent to which solutions of the stochastic integral equation (5.2) or of the differential equation (5.3) can be identified with solutions of the problem (5.1) with randomised stopping, and the existence and uniqueness of solutions to (5.2) by using fixed point theorem.

Lemma 10. *Suppose G is the solution to Problem 1. Then G also solves Problem 2 and vice versa.*

Proof. As discussed at the start of Section 5.3, since the payoff function is bounded by a linear function, so are w , h and the solution to Problem 1.

Let V be the value function of the randomised stopping problem at an instant when there is stopping opportunity. Then, we have

$$V(x) = \Gamma^{g,c}(x)g(x) + (1 - \Gamma^{g,c}(x))G(x) \quad (5.4)$$

Conditioning on the first event time T_1 of the Poisson process, and using the strong Markov property, we have

$$\begin{aligned}
G(x) &= \mathbb{E}^x[e^{-\beta\tau_c}g(X_{\tau_c})] \\
&= \mathbb{E}^x \left[\mathbb{E}^x[e^{-\beta\tau_c}g(X_{\tau_c})|T_1 = t] \right] \\
&= \int_0^\infty dt \lambda e^{-\lambda t} \mathbb{E}^x[e^{-\beta\tau_c}g(X_{\tau_c})|T_1 = t] \\
&= \mathbb{E}^x \left[\int_0^\infty dt \lambda e^{-\lambda t} \mathbb{E}^x[e^{-\beta\tau_c}g(X_{\tau_c})|T_1 = t, X_t] \right] \\
&= \mathbb{E}^x \left[\int_0^\infty dt \lambda e^{-\lambda t} e^{-\beta t} V(X_t) \right] \\
&= \mathbb{E}^x \left[\int_0^\infty dt \lambda e^{-\lambda t} e^{-\beta t} \{ \Gamma^{g,c}(X_t)g(X_t) + (1 - \Gamma^{g,c}(X_t))G(X_t) \} \right].
\end{aligned} \tag{5.5}$$

□

Lemma 11. *Suppose G solves Problem 2. Then G is C^∞ . Moreover, G solves Problem 3.*

Proof. It is a classical result (see for example Karatzas and Shreve (1991) Problem 4.3.1) that if $F : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is Borel measurable, and satisfies $\int_0^\infty e^{-a(\ln x)^2} F(x) d(\ln x) < \infty$ for some $a > 0$, then u^F is $C^\infty((0, \infty) \times (0, \infty))$ where $u^F(t, x)$ is defined by $u^F(t, x) = \mathbb{E}^x[F(X_t)] = \int_0^\infty F(y)P(t; x, y)dy$ and $P(t; x, y)$ is the transition density of a geometric Brownian motion.

Recall that G is of linear growth. Then V defined by (5.4), which is the weighted average of two functions of linear growth, is also of linear growth. In particular, u^V is C^∞ . Then by (5.5), $G(x) = \mathbb{E}^x \left[\int_0^\infty e^{-\beta t} \lambda e^{-\lambda t} V(X_t) dt \right] = \int_0^\infty \lambda e^{-(\beta+\lambda)t} u^V(t, x) dt$ is also C^∞ . Furthermore, we can obtain bounds on the derivatives of G , see for example the proof of Problem 4.3.1 in Karatzas and Shreve [31, p277]. For example, $\mathbb{E}^x[X^2|G''(X)|] < \infty$.

Now we show that G solves (5.3). We follow Pham [47, p43]. For $\delta > 0$, writing $t = s + \delta$ we have

$$\begin{aligned}
G(x) &= \mathbb{E}^x \left[\int_0^\delta e^{-\beta t} \lambda e^{-\lambda t} V(X_t) dt + \int_\delta^\infty e^{-\beta t} \lambda e^{-\lambda t} V(X_t) dt \right] \\
&= \mathbb{E}^x \left[\int_0^\delta \lambda e^{-(\beta+\lambda)t} V(X_t) dt \right] + \mathbb{E}^x \left[\mathbb{E} \left[\int_0^\infty \lambda e^{-(\beta+\lambda)(s+\delta)} V(X_{s+\delta}) ds \middle| \mathcal{F}_\delta \right] \right] \\
&= \mathbb{E}^x \left[\int_0^\delta \lambda e^{-(\beta+\lambda)t} V(X_t) dt \right] + \mathbb{E}^x \left[e^{-(\beta+\lambda)\delta} G(X_\delta) \right]
\end{aligned} \tag{5.6}$$

Let $\tau_n = \inf\{u : X_u \notin (\frac{x}{n}, nx)\}$. Since G is of class C^∞ , we apply Itô's formula to $e^{-(\beta+\lambda)t}G(X_t)$ to obtain

$$\begin{aligned} e^{-(\beta+\lambda)(\delta \wedge \tau_n)}G(X_{\delta \wedge \tau_n}) &= G(x) + \int_0^{\delta \wedge \tau_n} e^{-(\beta+\lambda)s}[\mathcal{L}G - (\beta + \lambda)G](X_s)ds \\ &\quad + \int_0^{\delta \wedge \tau_n} e^{-(\beta+\lambda)s}\sigma X_s G'(X_s)dW_s \end{aligned}$$

and hence

$$\mathbb{E}^x \left[e^{-(\beta+\lambda)(\delta \wedge \tau_n)}G(X_{\delta \wedge \tau_n}) \right] = G(x) + \mathbb{E}^x \left[\int_0^{\delta \wedge \tau_n} e^{-(\beta+\lambda)s}[\mathcal{L}G - (\beta + \lambda)G](X_s)ds \right].$$

By the linear growth condition of G , we obtain,

$$e^{-(\beta+\lambda)(\delta \wedge \tau_n)}G(X_{\delta \wedge \tau_n}) \leq e^{-(\beta+\lambda)(\delta \wedge \tau_n)}\kappa X_{\delta \wedge \tau_n} \leq \sup_{t \geq 0} e^{-(\beta+\lambda)t}\kappa X_t,$$

where κ is some positive constant. Also, since we have $\beta > \mu$, we obtain

$$\mathbb{E}^x \left[\sup_{t \geq 0} e^{-(\beta+\lambda)t}\kappa X_t \right] < \infty,$$

for geometric Brownian motion X with drift μ . Moreover, using the smoothness of G , we obtain

$$\mathbb{E}^x [|\mathcal{L}G - (\beta + \lambda)G(X_s)|] < \infty.$$

Letting n tend to infinity, by dominated and monotone convergence, we get

$$\mathbb{E}^x \left[e^{-(\beta+\lambda)\delta}G(X_\delta) \right] = G(x) + \mathbb{E}^x \left[\int_0^\delta e^{-(\beta+\lambda)s}[\mathcal{L}G - (\beta + \lambda)G](X_s)ds \right].$$

Plugging the above equation back into (5.6), we get

$$G(x) = \mathbb{E}^x \left[\int_0^\delta \lambda e^{-(\beta+\lambda)t}V(X_t)dt + G(x) + \int_0^\delta e^{-(\beta+\lambda)s}[\mathcal{L}G - (\beta + \lambda)G](X_s)ds \right]$$

and it follows that

$$0 = \mathbb{E}^x \left[\int_0^\delta e^{-(\beta+\lambda)s}[\mathcal{L}G - (\beta + \lambda)G + \lambda V](X_s)ds \right]. \quad (5.7)$$

Let $F(s) = \mathbb{E}^x[e^{-(\beta+\lambda)s}\{\mathcal{L}G - (\beta + \lambda)G + \lambda V\}(X_s)]$ and note that F is continuous on $[0, \infty)$. Dividing both sides of (5.7) by δ and sending δ to 0, we conclude from

the Mean-Value Theroem that there exists $\delta_n \downarrow 0$ such that $F(\delta_n) = 0$. Then, by continuity of F we conclude $F(0) = 0$, or equivalently $\mathcal{L}G - (\beta + \lambda)G + \lambda V = 0$. And by (5.4), we find

$$\mathcal{L}G - \beta G + \lambda \Gamma^{g,c}(g - G) = 0.$$

□

Lemma 12. *Suppose $f = f(x, h)$ is continuous and of at most linear growth, suppose $\epsilon > \beta$ and consider the ODE*

$$\mathcal{L}H(x) - \epsilon H(x) + f(x, H(x)) = 0. \quad (5.8)$$

Suppose H is a solution to (5.8) of at most linear growth. Then H has the probabilistic representation

$$H(x) = \mathbb{E}^x \left[\int_0^\infty e^{-\epsilon t} f(X_t, H(X_t)) dt \right]. \quad (5.9)$$

Proof. We have $H'' = \frac{2}{\sigma^2 x^2} \{-\mu x H' + \epsilon H - f(x, H)\}$ so that H is C^2 . Then, applying Itô's formula to $e^{-\epsilon t} H(X_t)$ we have

$$e^{-\epsilon(t \wedge \tau_n)} H(X_{t \wedge \tau_n}) = H(x) + \int_0^{t \wedge \tau_n} e^{-\epsilon s} [\mathcal{L}H(X_s) - \epsilon H(X_s)] ds + \int_0^{t \wedge \tau_n} e^{-\epsilon s} H'(X_s) \sigma X_s dW_s$$

where, as before $\tau_n := \inf\{u > 0 : X_u \notin (\frac{x}{n}, nx)\}$. Since the stopped stochastic integral is a martingale, taking expectations on both sides and using (5.8), we get

$$\mathbb{E}^x \left[e^{-\epsilon(t \wedge \tau_n)} H(X_{t \wedge \tau_n}) \right] = H(x) - \mathbb{E}^x \left[\int_0^{t \wedge \tau_n} e^{-\epsilon s} f(X_s, H(X_s)) ds \right].$$

Using the properties of exponential Brownian motion to conclude that $\mathbb{E}^x[\sup_{s \leq t} X_s] < Cx$ for some C constant, sending n to infinity, and using dominated convergence and the assumed linear growth of H ,

$$\mathbb{E}^x \left[e^{-\epsilon t} H(X_t) \right] = H(x) - \mathbb{E}^x \left[\int_0^t e^{-\epsilon s} f(X_s, H(X_s)) ds \right].$$

Then, since H is of linear growth and $\epsilon > \mu$, sending t to infinity we conclude

$$\begin{aligned} 0 &= \lim_{t \rightarrow \infty} \mathbb{E}^x [e^{-\epsilon t} H(X_t)] = H(x) - \lim_{t \rightarrow \infty} \mathbb{E}^x \left[\int_0^t e^{-\epsilon s} f(X_s, H(X_s)) ds \right] \\ &= H(x) - \mathbb{E}^x \left[\int_0^\infty e^{-\epsilon s} f(X_s, H(X_s)) ds \right]. \end{aligned}$$

Thus, H admits probabilistic representation (5.9). \square

Now, taking $\epsilon = \lambda + \beta$, $H = G$ and $f(x, h) = \lambda\{\Gamma^{g,c}(x)g(x) + (1 - \Gamma^{g,c}(x))h\}$, we conclude

$$G(x) = \mathbb{E}^x \left[\int_0^\infty dt \lambda e^{-\lambda t} e^{-\beta t} \{ \Gamma^{g,c}(X_t)g(X_t) + (1 - \Gamma^{g,c}(X_t))G(X_t) \} \right].$$

Proposition 6. *For each $c \in \{w, h, G\}$ there exists a unique G which has the probabilistic representation (5.2), is of class \mathcal{C}^2 and satisfies a linear growth condition.*

Proof. Denote by (M, d) the metric space

$$\begin{aligned} M &= \{f : (0, \infty) \mapsto (0, \infty), f \in C^2, 0 < f(x) < \kappa x \text{ for some } \kappa \in \mathbb{R}_+\}, \\ d(H_1, H_2) &= \sup_{x \in (0, \infty)} \left| \frac{H_1(x) - H_2(x)}{x} \right|. \end{aligned}$$

For c a perceived continuation value define $T^c : M \mapsto M$ by

$$T^c(F)(x) = \mathbb{E}^x \left[\int_0^\infty e^{-(\beta+\lambda)t} \lambda \frac{g(X_t)^2 + c(X_t)F(X_t)}{g(X_t) + c(X_t)} dt \right] > 0.$$

To see that $T^c(F) \in M$ note that $T^c(F)$ is of class \mathcal{C}^2 by Lemma 11, and since g and F are of linear growth and $\frac{g^2+cF}{g+c} \leq g + F$,

$$0 < T(F)(x) \leq \mathbb{E}^x \left[\int_0^\infty \lambda e^{-(\beta+\lambda)t} [F(X_t) + g(X_t)] dt \right] \leq \kappa x$$

where κ is some positive constant.

Next, we show that T^c is a contraction mapping. Then, by the Banach fixed point theorem there exists a unique function $m \in M$ such that $T^c(m) = m$. Thus there is a unique solution to Problem 2.

We show $d(T^c(H_1), T^c(H_2)) \leq \rho d(H_1, H_2)$ with $\rho < 1$. There are three cases to consider, namely $c = w$, $c = h^\lambda$ and $c = G$. We cover the first two cases together:

for $c = w$ and $c = h^\lambda$,

$$\begin{aligned}
|T(H_1)(x) - T(H_2)(x)| &\leq \mathbb{E}^x \left[\int_0^\infty e^{-(\beta+\lambda)t} \lambda \left| \frac{g^2 + cH_1}{g+c} - \frac{g^2 + cH_2}{g+c} \right| dt \right] \\
&= \int_0^\infty e^{-(\beta+\lambda)t} \lambda \mathbb{E}^x \left[\frac{c}{g+c} |H_1 - H_2| \right] dt \\
&\leq \int_0^\infty e^{-(\beta+\lambda)t} \lambda \mathbb{E}^x \left[X_t \left| \frac{H_1 - H_2}{X_t} \right| \right] dt \\
&\leq d(H_1, H_2) \lambda \int_0^\infty e^{-(\beta+\lambda)t} \mathbb{E}^x[X_t] dt = \rho d(H_1, H_2)x,
\end{aligned}$$

where $\rho = \frac{\lambda}{\lambda+\beta-\mu} < 1$. Similarly, when $c = G$, and using $0 < \frac{2g^2}{(g+H_1)(g+H_2)} < 2$ for the last line

$$\begin{aligned}
&|T(H_1)(x) - T(H_2)(x)| \\
&\leq \mathbb{E}^x \left[\int_0^\infty e^{-(\beta+\lambda)t} \lambda \left| \frac{g^2 + H_1^2}{g+H_1} - \frac{g^2 + H_2^2}{g+H_2} \right| dt \right] \\
&= \int_0^\infty e^{-(\beta+\lambda)t} \lambda \mathbb{E}^x \left[\left| g + H_1 - \frac{2gH_1}{H_1+g} - g - H_2 - \frac{2gH_2}{g+H_2} \right| \right] dt \\
&= \int_0^\infty e^{-(\beta+\lambda)t} \lambda \mathbb{E}^x \left[X_t \left| \frac{H_1 - H_2}{X_t} \right| \left| 1 - \frac{2g^2}{(g+H_1)(g+H_2)} \right| \right] dt \\
&\leq d(H_1, H_2) \lambda \int_0^\infty e^{-(\beta+\lambda)t} \mathbb{E}^x[X_t] dt \\
&= \rho x d(H_1, H_2)
\end{aligned}$$

as required. □

Theorem 9. Suppose $\Gamma^{g,c} = \frac{g}{g+c}$. Then the solution to any one of the three formulations is the unique solution to all of them.

Proof. Theorem 9 follows immediately from the following results

1. If f^{STF} is the solution of the stopping problem formulation then f^{STF} solves (5.2).
2. If f^{SFF} is of polynomial growth and solves (5.2) then it also solves (5.3).
3. If f^{ODEF} is of linear growth and solves (5.3) then it also solves (5.2).
4. There is a unique solution to the Stochastic Formulation with Feedback problem.

□

Note that there will be solutions of (5.2) and (5.3) which are not of linear growth. These solutions might be identified with bubbles in the sense of Scheinkman and Xiong (2003). They correspond to solutions of Problems 2 and 3 which involve internally consistent valuations where the agent's current over-valuation of the solution is justified by an overvaluation at future candidate stopping times also. However, they do not have a representation as a solution of the stopping time formulation. We will not be concerned with such solutions.

5.4 Linear payoffs

In this section we suppose $g(x) = x$. Then in the classical optimal stopping problem it is always optimal to exercise immediately, and $w(x) = x$. For the problem in which exercise times are restricted to event times, it's optimal to stop at the first Poisson event time and receive the payoff, that is

$$\begin{aligned} h(x) &= \mathbb{E}^x \left[\int_0^\infty \lambda e^{-\lambda t} e^{-\beta t} g(X_t) dt \right] = \int_0^\infty dt \lambda e^{-\lambda t} e^{-\beta t} \mathbb{E}^x [X_t] \\ &= \int_0^\infty dt \lambda e^{-\lambda t} e^{-\beta t} x e^{\mu t} = \rho x \end{aligned}$$

where $\rho = \frac{\lambda}{\lambda + \beta - \mu} \in (0, 1)$. There are three possible forms for the value of the randomised stopping problem depending on which version of the perceived continuation value we use. Using $\Gamma^{g,c} = \frac{g}{g+c}$, (5.3) can be rewritten as

$$\mathcal{L}G - (\beta + \lambda)G + \lambda \frac{g^2 + cG}{g + c} = 0. \quad (5.10)$$

5.4.1 Perceived continuation value $c = w$

In this case, we denote by G^w the continuation value function, then we find from (5.10) that $G^w(x) = \psi^w x$ where ψ^w solves

$$\mu \psi x - (\beta + \lambda) \psi x + \lambda \frac{(1 + \psi)x^2}{2x} = 0.$$

We find $\psi^w = \frac{\rho}{2-\rho}$.

5.4.2 Perceived continuation value $c = h$

In this case, we denote by G^h the continuation value function, then $G^h(x) = \psi^h x$ where ψ^h solves

$$\mu\psi x - (\beta + \lambda)\psi x + \lambda \frac{(1 + \rho\psi)x^2}{(1 + \rho)x} = 0.$$

We find $\psi^h = \frac{\rho}{1 + \rho - \rho^2}$.

5.4.3 Perceived continuation value $c = G$

In this case, the perceived continuation value c coincides with the continuation value G . We simply use $G = G^G$ as the continuation value function, then $G(x) = \psi^G x$ where ψ^G solves

$$\mu\psi x - (\beta + \lambda)\psi x + \lambda \frac{(1 + \psi^2)x^2}{(1 + \psi)x} = 0 \quad (5.11)$$

We find $\psi^G = \sqrt{\frac{1}{4(1-\rho)^2} + \frac{\rho}{1-\rho}} - \frac{1}{2(1-\rho)}$, where we take the larger root of (5.23) as this root lies in $(0, 1)$.

5.4.4 Discussion

We will explain in the discussion why $\rho > \psi^G > \psi^h > \psi^w$, and this is confirmed graphically in Figure 5.1.

First observe that as $\mathcal{T}_\lambda \subset \mathcal{T}$ we must have $h \leq w$, and since h is optimal for stopping at event times of the Poisson process we must have $\max\{G^w, G^h, G\} < h$.

In the problem with a linear payoff it is always optimal to stop as soon as possible both in the classical optimal stopping problem, and in the stopping problem in which stopping times are restricted to be event times of the Poisson process. This remains true in the randomised stopping problem, to the extent that the problem value is maximised if the probability of stopping is maximised. Since the probability of stopping $\frac{g}{g+c}$ is maximised when c is minimised, it follows from the inequalities $G < h < w$ that the value functions have order $G > G^h > G^w$. Hence, $\psi^G > \psi^h > \psi^w$. Further, all these valuations are dominated by the case of optimal stopping where stopping times are constrained to be event times of the Poisson process, and so $\psi^G < \rho$.

For all specifications of continuation value, ψ^c has limiting values $\psi^c(0+) = 0$ and $\psi^c(1-) = 1$. When λ is very small, T_1 is likely to be large, $e^{-\beta T_1} X_{T_1}$ is small with large probability, and the value function is small. Conversely, if $\beta - \mu$ is small, $\mathbb{E}[e^{-\beta T_k^\lambda} X_{T_k^\lambda}]$ is close to unity. Although, the agent would benefit most from stopping at each and every opportunity, the losses from not stopping are not great.

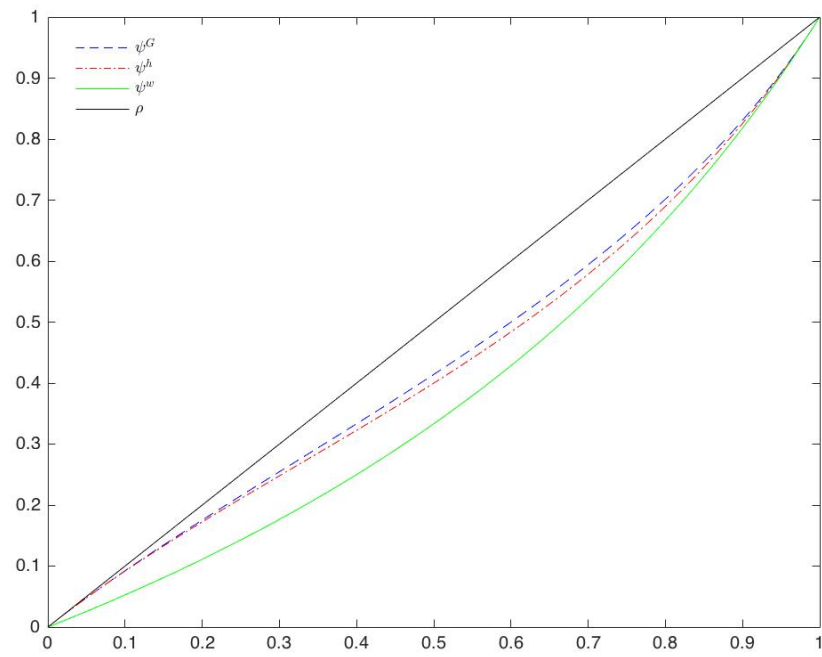


Figure 5.1: A plot of ψ^w , ψ^h and ψ^G as functions of ρ , as well as the line $y(\rho) = \rho$. Note that for $g(x) = x$ we have $h(x) = \rho x$ and $G^c(x) = \psi^c x$.

Note that ψ^c is increasing in ρ for each $c \in \{w, h, G\}$. This corresponds to the value function being increasing in λ . Consider first the case $c = w$. As λ increases, there are more chances to stop. Since w does not depend on λ , the probability of stopping, conditional on an opportunity to stop, does not depend on λ . Hence, a simple coupling argument gives that as λ increases the stopping time gets smaller and therefore the value function increases. Now consider the case $c = h$. As λ increases, there are more opportunities to stop. However, h is increasing in λ , and so at each opportunity to stop the agent is less likely to stop. This second factor is less significant than the first, and overall the rate of stopping $\lambda\Gamma(x, h)$ goes up. Hence ψ^h is increasing in λ . Finally suppose $c = G$. Again, increasing λ increases the stopping opportunities which has the impact of increasing the value function. However, this reduces the probability of stopping, which has the effect of reducing the size of any increase in value function, but not to the extent of preventing overall increases.

5.5 Call payoffs

Our goal in this section is to move beyond linear payoffs to call payoffs. In particular we will assume $g(x) = (x - K)^+$. By a scaling argument it is possible to reduce the case of general strike to unit strike, and in all our numerical examples we will assume $K = 1$, but for the present we allow general K . Section 2 provides explicit formulae for w in (2.8).

5.5.1 Explicit formula for h

We can solve for h by noting that it is optimal to stop at (t, X_t) if and only if there is an event of the Poisson process and $h(X_t) \leq (X_t - K)$. We expect that there is a critical value L^λ such that it is optimal to stop at (t, X_t) if and only if $X_t > L^\lambda$. Then we have $\mathcal{L}h - \beta h = 0$ for $x \leq L^\lambda$ and $\mathcal{L}h - (\beta + \lambda)h + \lambda g = 0$ for $x \geq L^\lambda$ (See Dupis and Wang (2005)). We have value matching and smooth fit at $x = L^\lambda$, and from the fact that L^λ separates the stopping and continuation regions, we have $h(L^\lambda) = g(L^\lambda) = (L^\lambda - K)^+$. We find

$$h(x) = \begin{cases} Cx^\theta, & x < L^\lambda \\ \rho x - \frac{\lambda K}{\lambda + \beta} + C_1 x^\gamma, & x \geq L^\lambda \end{cases} \quad (5.12)$$

where

$$\gamma = \left(\frac{1}{2} - \frac{\mu}{\sigma^2} \right) - \sqrt{\left(\frac{1}{2} - \frac{\mu}{\sigma^2} \right)^2 + \frac{2(\beta + \lambda)}{\sigma^2}} < 0, \quad (5.13)$$

and $L^\lambda = \frac{\lambda + \beta - \mu}{\lambda + \beta} \frac{\gamma}{\gamma - 1} \frac{\theta}{\theta - 1} K$, $C = \frac{\lambda}{\lambda + \beta} \frac{\gamma}{\gamma - \theta} \frac{1}{\theta - 1} K (L^\lambda)^{-\theta}$, and $C_1 = \frac{\lambda}{\lambda + \beta} \frac{\theta}{\gamma - \theta} \frac{1}{\gamma - 1} K (L^\lambda)^{-\gamma}$.

Note that $\lim_{\lambda \uparrow \infty} \gamma = -\infty$ and hence $\lim_{\lambda \uparrow \infty} L^\lambda = L^*$. Note further that $\lim_{\lambda \uparrow \infty} C = \frac{1}{\theta - 1} K (L^*)^{-\theta} = \frac{1}{\theta} (L^*)^{1 - \theta}$ where we use $L^* = \frac{\theta}{\theta - 1} K$ and similarly $\lim_{\lambda \uparrow \infty} C_1 = 0$, and moreover $C_1 x^\gamma \rightarrow 0$ for fixed x . Hence $\lim_{\lambda \uparrow \infty} h(x) = w(x)$.

5.5.2 Perceived continuation value $c = w$

The first randomised stopping problem we consider is for the case where the continuation value is the value of the problem with no restrictions on the exercise time. By (5.3), we have

$$\mathcal{L}G^w - (\beta + \lambda)G^w + \lambda \frac{g^2 + wG^w}{g + w} = 0, \quad x \in (0, \infty), \quad (5.14)$$

For $x \in (0, K)$, we have $g(x) = 0$ and thus G^w satisfies

$$\mathcal{L}G^w - \beta G^w = 0, \quad x \in (0, K). \quad (5.15)$$

Recall that $L^* = \frac{\theta}{\theta - 1} K$. For $x \in [L^*, \infty)$, we have $w(x) = g(x)$ and thus G^w satisfies

$$\mathcal{L}G^w - (\beta + \frac{\lambda}{2})G^w + \frac{\lambda}{2}g = 0, \quad x \in [L^*, \infty). \quad (5.16)$$

The general solution to (5.15) is

$$G^w(x) = Bx^\theta + B_0x^{\theta_0}$$

where θ and θ_0 are given by (2.4) and (2.5) respectively. From the boundary condition $G^w(0+) = 0$ we must have $B_0 = 0$.

Similarly, the general solution to $\mathcal{L}G^w - (\beta + \frac{\lambda}{2})G^w = 0$ is given by $G^w(x) = B_3x^{\alpha_+} + B_4x^{\alpha_-}$ where $\alpha_+ > 1$ and $\alpha_- < 0$ are given by $\alpha_\pm = (\frac{1}{2} - \frac{\mu}{\sigma^2}) \pm \sqrt{(\frac{1}{2} - \frac{\mu}{\sigma^2})^2 + \frac{2\beta + \lambda}{\sigma^2}}$.

A particular solution to (5.16) is given by

$$G^w(x) = \frac{\lambda}{\lambda + 2\beta - 2\mu}x - \frac{\lambda K}{\lambda + 2\beta} = \frac{\rho}{2 - \rho}x - \frac{\lambda K}{\lambda + 2\beta}$$

Since the solution G^w we are looking for is of linear growth rate, we require $B_3 = 0$

and it follows that for $x \in (L^*, \infty)$

$$G^w(x) = \psi^w x - \frac{\lambda}{\lambda + 2\beta} K + B_4 x^{\alpha-}, \quad (5.17)$$

for a constant B_4 to be determined.

The goal is to construct a C^2 solution for G^w on $(0, \infty)$. Fix a solution for G^w on $(0, K)$ by fixing B_1 . We can use value matching and smooth fit at K to give values for G^w and its first derivative at K and hence to construct a (numerical) solution to (5.14) on $[K, L^*)$. Value matching at L^* can be used to construct a solution to (5.16) on (L^*, ∞) , and in particular to fix B_4 in (5.17). In general there will be no first order smooth fit at L^* . However, by adjusting B_1 we can construct a solution which is C^1 at K and L^* and hence C^1 on $(0, \infty)$. This is the solution we want.

Note that if we set $g = 0$ at $x = K$ then (5.14) reduces to (5.15), and if we set $g = w$ at $x = L^*$ then (5.14) reduces to (5.16). As a result, if we have a solution which is C^1 at K and L^* then the second derivatives also match at these points, and our C^1 solution is actually C^2 .

5.5.3 Perceived continuation value $c = h$

Now suppose we take as the continuation value the value of the game under optimal stopping when the stopping opportunities are the event times of a Poisson Process, rate λ . We have that G^h satisfies

$$\mathcal{L}G^h - (\beta + \lambda)G^h + \lambda \frac{g^2 + hG^h}{g + h} = 0 \quad x \in (0, \infty), \quad (5.18)$$

where h is given by (5.12). Note that g changes form at K and h changes form at L^λ so that (5.18) can usefully be split into three regions. As in the previous case, the boundary condition at $0+$ is such that the solution on $(0, K]$ takes the form $G^h(x) = Dx^\theta$ for some constant D . Temporarily fixing D , value matching and first-order smooth fit at K allows us to construct a solution on $[K, \infty)$. We want the solution for which $\lim_{x \rightarrow L^\lambda} \frac{G^h(x)}{x} = \psi^h$; we adjust D until this is the case. Again, since g and h are continuous at K and L^λ , the C^1 solution from (5.18) is automatically C^2 .

5.5.4 Perceived continuation value $c = G$

We distinguish between the two regions for (5.3),

$$\mathcal{L}G - \beta G = 0, \quad x \in (0, K) \quad (5.19)$$

$$\mathcal{L}G - (\beta + \lambda)G + \lambda \frac{g^2 + G^2}{g + G} = 0, \quad x \in [K, \infty). \quad (5.20)$$

The general solution to (5.19) on $(0, K)$ is given by $G(x) = Ex^\theta$ for some constant E . Fixing E and using value matching and first order smooth fit we can construct (numerically) a C^1 solution for G on $(0, \infty)$. Finally, we can adjust E until we obtain a solution with linear growth which satisfies $\lim_{x \rightarrow \infty} \frac{G(x)}{x} = \psi^G$.

5.5.5 Comparison of the different solutions

Figure 6.14 plots the various value functions w , h , G^w , G^h and G together with the payoff $g(x) = (x - 1)^+$. w is the largest of the value functions, reflecting the fact that stopping is unrestricted, and that stopping is optimal. Next largest is h which involves optimal stopping from the event times of the Poisson process: optimality means that $h \geq \max\{G^w, G^h, G\}$.

Since $w > h$, when we compare the stopping probability for randomised stopping under continuation value w compared with that of h we expect to stop less frequently. In general, discounting means that above and not too close to the strike it is beneficial to stop sooner. Hence $G^h > G^w$. (Below the strike $g \equiv 0$, and the probability of stopping is zero. Just above the strike, stopping is more common for $c = h$ than for $c = w$, and stopping is sub-optimal in this case; nonetheless, this regime is small and $G^h > G^w$.)

Similar reasoning justifies why $G < h$ leads to $G > G^h$. From Figure 6.14 we see that $h - G \ll w - h$ and from this we expect that $G - G^h \ll G^h - G^w$, where by \ll we mean *much smaller than* in a qualitative sense. Again the evidence from Figure 6.14 supports this conclusion.

Figure 5.3 shows the impact of increased stopping opportunities and shows the value function as a function of x for various values of λ . Surprisingly, in general the value function is non-monotonic in λ . For large values of x (see panel (a)) we have that $G(x)$ is monotonic in λ : for large x it is always optimal to stop and hence more stopping opportunities are beneficial (recall that asymptotically $\frac{G(x)}{x} \rightarrow \psi^G$ and ψ^G is monotonic in λ , Figure 5.1). However, this monotonicity does not propagate to all values of x . For x close to the strike (see Panel (b)) the value function is non-monotonic. This reflects the multiple impacts of increasing λ ; it increases the

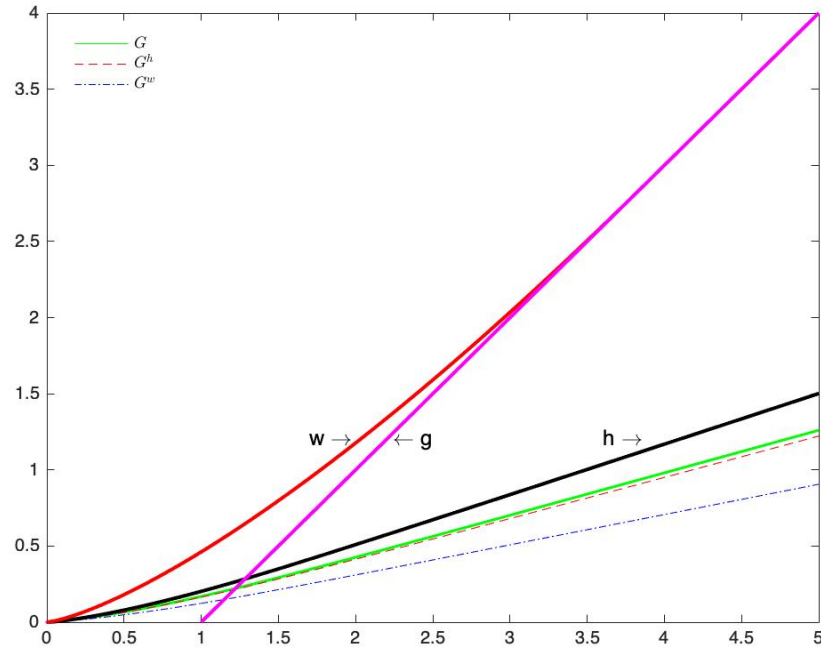


Figure 5.2: The value functions depicted are based the parameter set : $(\beta, \mu, \sigma, K, \lambda) = (5, 3, 2, 1, 1)$; $w > h > G > G^h > G^w$ always holds.

stopping opportunities and hence also the rate of stopping. Near the strike, since stopping is worse than continuing, more stopping can reduce the value function. Overall, the impact of increasing the rate stopping opportunities is ambiguous.

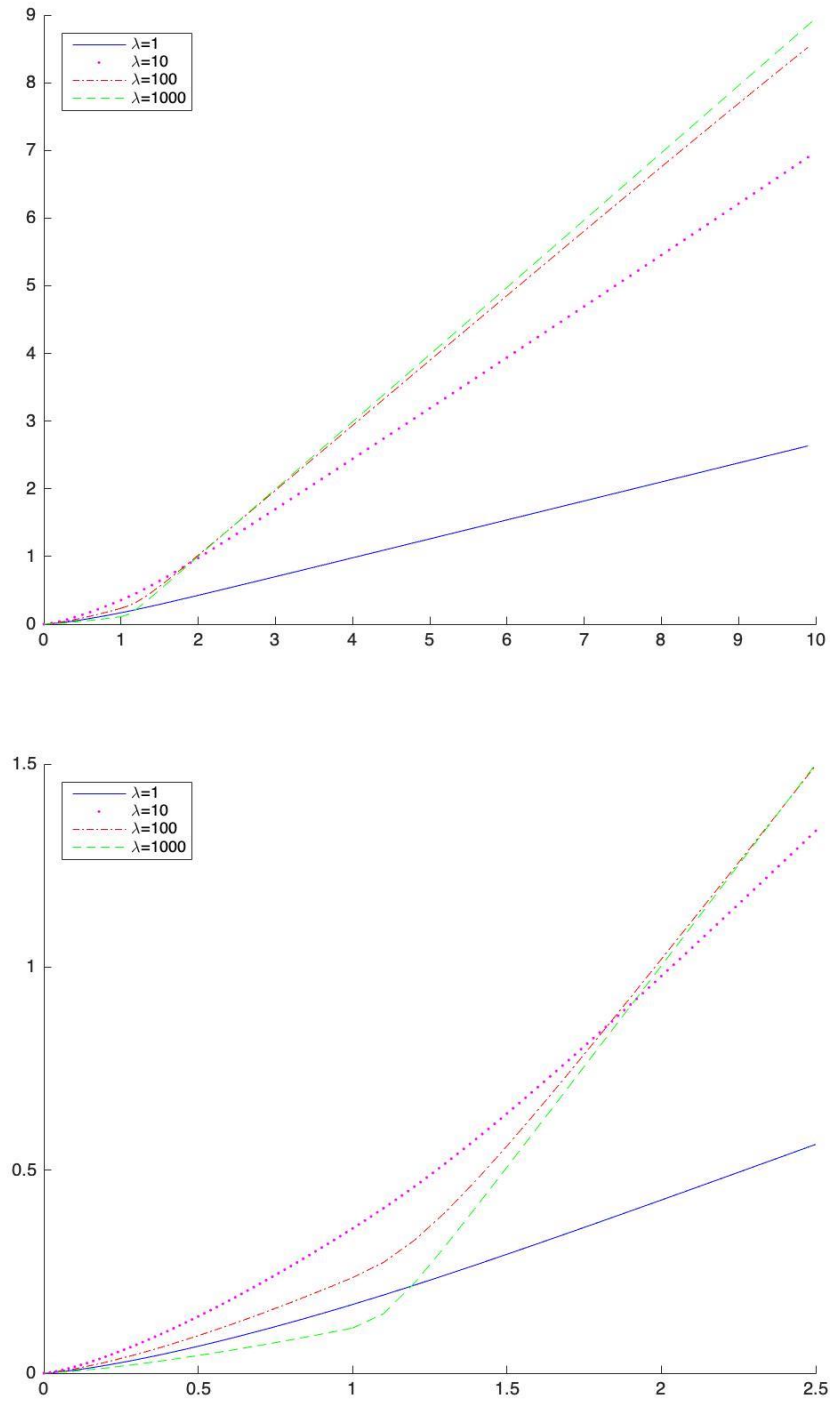


Figure 5.3: In the top plot, the value functions are seen to be increasing in λ for large x . In the bottom plot, we see that this monotonicity does not hold for λ near the strike. Other parameters are $(\beta, \mu, \sigma, K) = (5, 3, 2, 1)$

5.5.6 Realised rate of stopping

The probability of stopping at each price level of X is expressed using Γ , and is plotted in Figure 5.4.

The vertical dash-dot lines in Figure 5.4 represent the critical thresholds L^λ which divide the continuation regions from the stopping regions in the optimal stopping problem in which stopping can only occur at event times of the Poisson process. We see that these critical thresholds are increasing in λ , and that they increase to L^* (represented by the vertical dash-dot line labelled $\lambda = \infty$). The curves represent the probability of stopping $\Gamma^{g,G}(x) = \frac{g(x)}{g(x)+G(x)}$.

As we can see from Figure 5.4, for fixed λ , the probability of stopping is increasing in x and converges to a constant. For large x , we have $\Gamma \geq 0.5$ since $g \geq h \geq G$. Moreover, we can see from Figure 5.4 that Γ is monotone decreasing in λ when x is large but loses the monotonicity when x is smaller. Again, this reflects the non-monotonicity of G in λ .

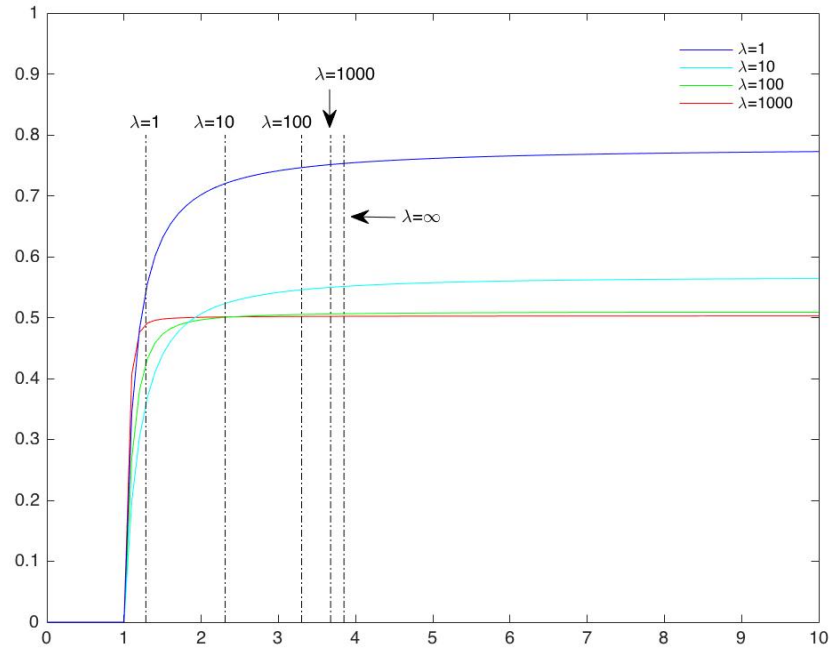


Figure 5.4: The stopping probability as a function of x , for varying λ . $(\beta, \mu, \sigma, K) = (5, 3, 2, 1)$; The vertical dash-dot represents the optimal stopping level for h when $\lambda = 1, \lambda = 10, \lambda = 100$ and $\lambda = 1000$ respectively (from left to right.); $\Gamma^{g,G} = \frac{g}{g+G}$

5.6 Towards a model of continuous stopping

5.6.1 Modification of the randomising stopping rule

If we assume that the probability of stopping (conditional on an event of the Poisson process) is a constant $p > 0$, independent of X_t , (which is the case when the payoff is linear or equivalently when the strike price K is 0), then the time of stopping is an exponentially distributed random variable with rate $p\lambda$. Then, as opportunities to stop come faster and faster ($\lambda \rightarrow \infty$), the time of stopping converges to 0, almost surely. Without modification to our model, if stopping opportunities become more and more frequent, then in the limit the randomising stopping rule will be degenerate and will involve stopping immediately wherever $g > 0$.

In order to avoid this degenerate limit we consider biasing the continuation probability towards continuing: we modify the stopping probability (previously $\Gamma(g, c) = \frac{g}{g+c}$) to

$$\Gamma_\xi = \Gamma_\xi(g, c) = \frac{g}{g + \xi c},$$

where ξ is a positive constant which we will specify in Section 5.6.2. As in Section 5.4, in the case of linear payoffs we can derive exact expressions for the value function: these take the form $V_\xi^c(x) = \psi_{\lambda, \xi}^c x$. For $c = w$, $\psi_{\lambda, \xi}^w$ solves

$$\mu\psi x - (\beta + \lambda)\psi x + \lambda \frac{(1 + \xi\psi)}{1 + \xi} x = 0,$$

we find

$$\psi_{\lambda, \xi}^w = \frac{\rho}{1 + \xi - \xi\rho}. \quad (5.21)$$

For $c = h$, $\psi_{\lambda, \xi}^h$ solves

$$\mu\psi x - (\beta + \lambda)\psi x + \lambda \frac{1 + \xi\rho\psi}{1 + \xi\rho} x = 0,$$

we find

$$\psi_{\lambda, \xi}^h = \frac{\rho}{1 + \xi\rho - \xi\rho^2}, \quad (5.22)$$

For $c = G$, $\psi_{\lambda, \xi}^G$ solves

$$\mu\psi x - (\beta + \lambda)\psi x + \lambda \frac{1 + \xi\psi^2}{1 + \xi\psi} x = 0 \quad (5.23)$$

we find

$$\psi_{\lambda,\xi}^G = -\frac{1}{2\xi(1-\rho)} + \sqrt{\frac{1}{4\xi^2(1-\rho)^2} + \frac{\rho}{\xi(1-\rho)}}. \quad (5.24)$$

Figure 5.5 shows the impact of varying ξ . We can see that the values of linear payoffs are decreasing in ξ . Increasing ξ decreases the probability of stopping for all cases, and since stopping is optimal everywhere, discounting reduces the value of the payoff. Hence $\psi_{\lambda,\xi}^c$ is decreasing in ξ for $c \in \{w, h, G\}$. Moreover, since $G < h < w$ we find $\psi_{\lambda,\xi}^G > \psi_{\lambda,\xi}^h > \psi_{\lambda,\xi}^w$.

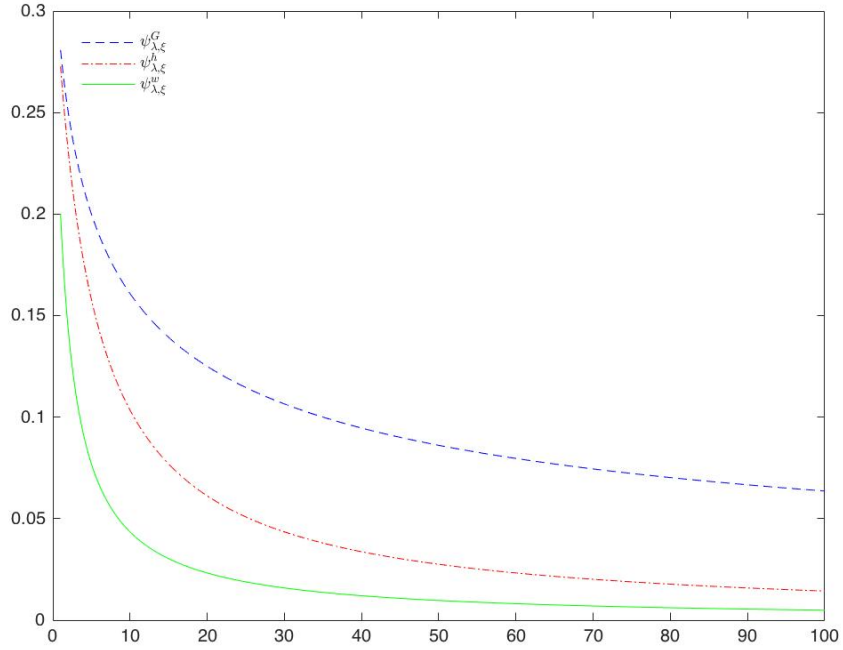


Figure 5.5: $(\beta, \mu, \sigma, \lambda) = (5, 3, 2, 1)$: plots of $\psi_{\lambda,\xi}^G$, $\psi_{\lambda,\xi}^h$ and $\psi_{\lambda,\xi}^w$ as functions of ξ .

5.6.2 Making ξ dependent on λ

Now we consider the impact of varying λ and ξ in a systematic manner. Suppose $c(x) = \kappa g(x)$ for some constant κ (for example, if $g(x) = x$ we find $c(x) = \kappa x$ for some κ .) Then $\Gamma_\xi(g, c) = \frac{1}{1+\xi\kappa}$ is independent of x , and the rate of stopping is $\frac{\lambda}{1+\kappa\xi}$. We want to choose $\lambda \uparrow \infty$, $\xi \uparrow \infty$ in such a way that the rate of stopping converges to a non-trivial rate. In particular we want to choose $\xi = \xi(\lambda)$ such that $\lim_{\lambda \uparrow \infty} \frac{\lambda}{1+\kappa\xi(\lambda)}$ exists in $(0, \infty)$. Then, as opportunities to stop (from the Poisson process) become universal, the probability of stopping (in a fixed and finite time

interval $[0, T_\epsilon]$) converges to a probability in $(0, 1)$.

Motivated by this heuristic we take $\xi = \frac{\lambda}{\eta}$ for $\eta \in (0, \infty)$. Then $\lambda\Gamma(g, c) = \frac{\eta\lambda g}{\eta g + \lambda c}$. In Figure 5.6 we plot $\psi_{\lambda, \frac{\lambda}{\eta}}^c$ as a function of η for $c \in \{w, h, G\}$. We see that as η increases $\psi_{\lambda, \lambda/\eta}^c$ increases. Moreover, $\psi_{\lambda, \lambda/\eta}^G > \psi_{\lambda, \lambda/\eta}^h > \psi_{\lambda, \lambda/\eta}^w$ and the first two are almost indistinguishable for even moderately large values of η .

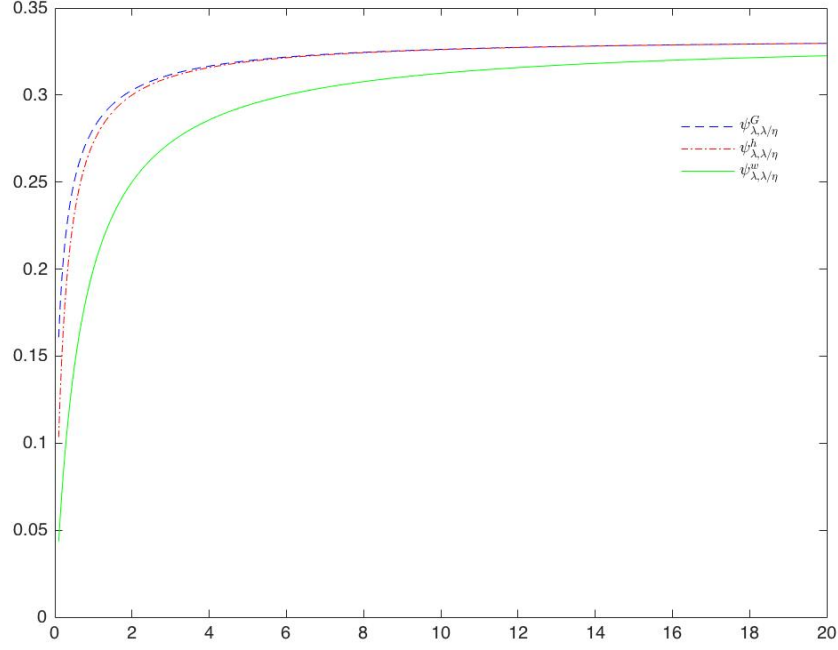


Figure 5.6: $\psi_{\lambda, \lambda/\eta}^c$ as a function of η . $(\beta, \mu, \sigma, \lambda) = (5, 3, 2, 1)$.

Our main interest is in fixing η and letting both λ and $\xi = \frac{\lambda}{\eta}$ get large. The values of ψ^c are plotted as functions of λ in Figure 5.7. Again we see $\psi_{\lambda, \lambda/\eta}^G > \psi_{\lambda, \lambda/\eta}^h > \psi_{\lambda, \lambda/\eta}^w$. We also have that $\psi_{\lambda, \frac{\lambda}{\eta}}^h$ and $\psi_{\lambda, \frac{\lambda}{\eta}}^w$ converge to the same limit. This is because, as λ increases to infinity h^λ converges to w and so the continuation value is the same for these two specifications. However, this is a limiting result, and when λ is small or moderate, $\psi_{\lambda, \lambda/\eta}^h$ is closer to $\psi_{\lambda, \lambda/\eta}^G$ than $\psi_{\lambda, \lambda/\eta}^w$, recovering the result of Section 5.5.5.

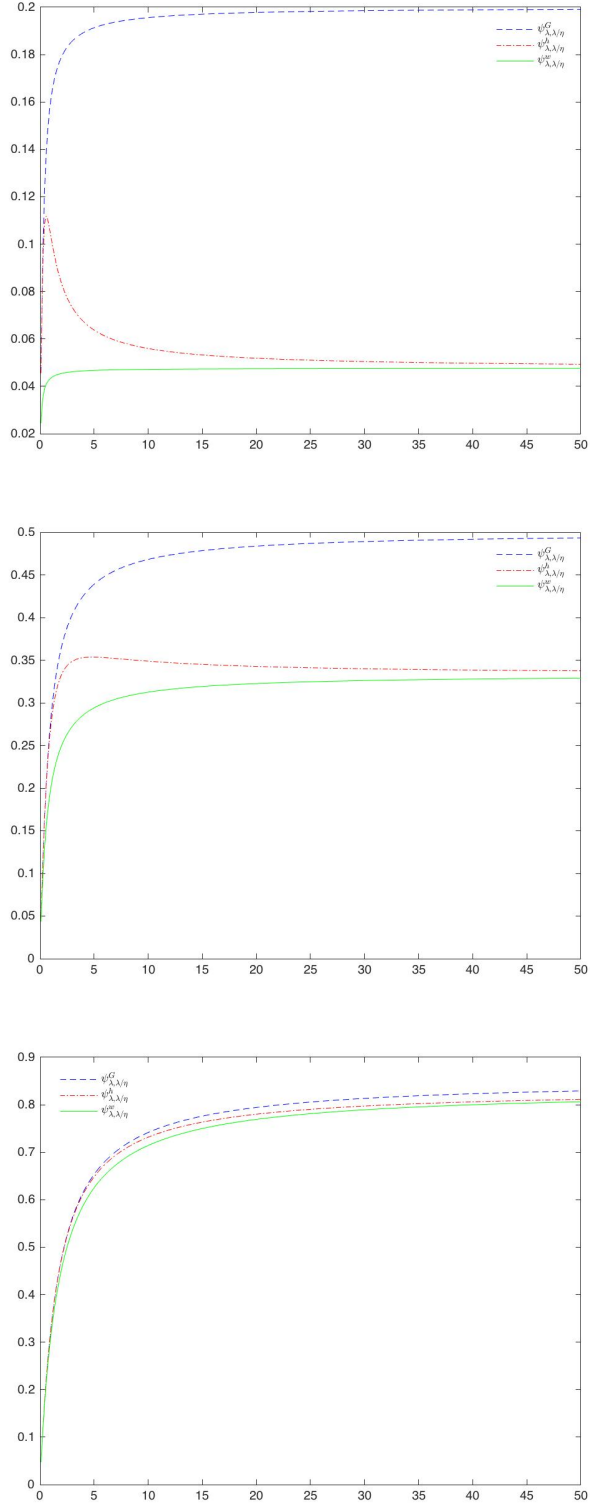


Figure 5.7: Under the new rule Γ_ξ for $\xi = \lambda/\eta$, we find $\lim_{\lambda \uparrow \infty} \psi_{\lambda, \lambda/\eta}^G < 1$. From top to bottom, $\eta = 0.1, 1, 10$ respectively

Recall the definitions of $\psi_{\lambda,\xi}^c$ in (5.24) - (5.21) and consider $\lim_{\lambda \uparrow \infty} \psi_{\lambda,\lambda/\eta}^c$. Define

$$k_w^* = \lim_{\lambda \uparrow \infty} \psi_{\lambda,\lambda/\eta}^w = \frac{\eta}{\eta + \beta - \mu} \quad (5.25)$$

$$k_h^* = \lim_{\lambda \uparrow \infty} \psi_{\lambda,\lambda/\eta}^h = \frac{\eta}{\eta + \beta - \mu} \quad (5.26)$$

$$k_G^* = \lim_{\lambda \uparrow \infty} \psi_{\lambda,\lambda/\eta}^G = -\frac{\eta}{2(\beta - \mu)} + \sqrt{\frac{\eta^2}{4(\beta - \mu)^2} + \frac{\eta}{\beta - \mu}} \quad (5.27)$$

Then k_c^* describes the value function (in the limit of large λ) for linear payoffs in the sense that for $g(x) = x$, $\lim_{\lambda \uparrow \infty} V_{\lambda,\lambda/\eta}^c(x) = k_c^*x$. By letting λ and ξ tend to infinity simultaneously we have obtained a non-degenerate limit. The limiting case $\lambda = \infty$ corresponds to a continuous flow of stopping opportunities, but with a non-trivial probability of stopping in each fixed interval $[0, T]$. In particular, $G^\lambda = G^{\lambda,\xi=\lambda/\eta,c}$ solves $0 = \{\mathcal{L}G^\lambda - \beta G^\lambda + \lambda \Gamma_{\lambda/\eta}(g, c)(g - G^\lambda)\} = \mathcal{L}G^\lambda - \beta G^\lambda + \lambda \frac{\eta g(g - G^\lambda)}{\eta g + \lambda c}$. Assuming $G_\eta^c = \lim_{\lambda \uparrow \infty} G^{\lambda,\xi=\lambda/\eta,c}$ exists and that we can swap the order of taking limits and differentiation we obtain that G_η^c solves

$$\mathcal{L}G_\eta^c - \beta G_\eta^c + \eta \frac{g}{c} (g - G_\eta^c) = 0. \quad (5.28)$$

For the case where the strike price is 0 (i.e. $g(x) = x$), the above ODE can be solved analytically and the solution is given by $G_\eta^c(x) = k_c^*x$ with k_c^* given by (5.25)-(5.27). Figure 5.8 shows the convergence of $G_{\eta,\lambda}$ by sending λ to infinity.

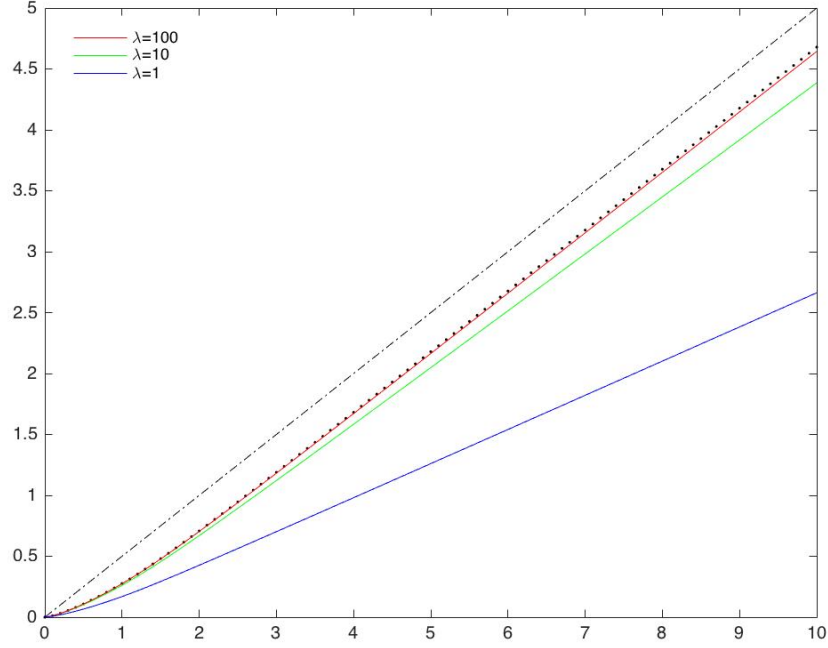


Figure 5.8: $(\beta, \mu, \sigma, K, \eta) = (5, 3, 2, 1, 1)$; the dash-dot line represents G^η ; the black-dot is the numerical solution to (5.28) in the case $c = G$; we can see that the dash-dot line and the black-dot line are parallel; $G_{\eta, \lambda}$ is increasing in λ and the slopes of $G_{\eta, \lambda}$ converges to that of G_η when λ gets large

5.6.3 Alternative formulation of the limiting case

In this section we propose a problem in continuous time in which the value function solves the same equation as that derived in the previous section, and hence represents a candidate continuous-time randomised stopping problem.

Suppose stopping opportunities occur as events of a time-inhomogeneous Poisson process with rate $\Lambda_\eta^c = \Lambda_\eta^c(x)$ where

$$\Lambda_\eta^c(x) = \frac{\eta g(x)}{c(x)}. \quad (5.29)$$

and that the option is exercised at *every* stopping opportunity. Here, as always, c is the continuation value, and in this model the rate of stopping depends on the ratio of the instantaneous payoff to the continuation value. Note that we identify stopping opportunities via an inhomogeneous Poisson process rather than by thinning a homogeneous Poisson process of rate λ . But h is the value function under the

assumption that there is a homogeneous Poisson process. Hence it does not makes much sense to consider $c = h$. Therefore, we will focus the two cases of $c \in \{w, G\}$.

The expected discounted reward from stopping can be represented via the stochastic formulation

$$G_\eta^c(x) = \mathbb{E}^x \left[\int_0^\infty \Lambda_\eta^c(X_t) e^{-\int_0^t \Lambda_\eta^c(X_s) ds} e^{-\beta t} g(X_t) dt \right]. \quad (5.30)$$

By analogy with the results in the previous section we assume that (5.30) has a unique solution, and that this solution is the unique solution of linear growth to the the ordinary differential equation

$$\mathcal{L}G(x) - [\beta + \Lambda(x)]G(x) + \Lambda(x)g(x) = 0. \quad (5.31)$$

Substituing for Λ in (5.31) we find that G solves (5.28). (This justifies why we have used the same notation $G = G_\eta^c$ for the value function in both Section 5.6.2 and in this section.) Thus, we have another interpretation for the continuous case ($\lambda \rightarrow \infty$) under the biased randomising stopping rule $\Gamma_{\xi=\lambda/\eta}$. This agent is employing a strategy of stopping at the first event time of an inhomogeneous Poisson process with rate $\Lambda_\eta^c(X_t) = \frac{\eta g(X_t)}{c(X_t)}$.

Linear payoffs

If $g(x) = x$ then it is always optimal to exercise immediately and $w(x) = x$. Then, in the case $c = w$ it follows from trying the candidate $G(x) = kx$ in (5.31) that $G_\eta^w(x) = k_w^* x$ where k_w^* is given by (5.25). Similarly, in the case $c = G$ we find $G_\eta^G(x) = k_G^* x$ where k_G^* is given by (5.27).

Since $G_\eta^G(x) < w(x)$ we find $\Lambda_\eta^G(\cdot) > \Lambda_\eta^w(\cdot)$ and hence when the continuation value is given by G we stop sooner than when the continuation value is given by w . This explains why $G^G > G^w$, or equivalently $k_G^* > k_w^*$.

Call payoffs

Now we suppose $g(x) = (x - 1)^+$ and consider numerical solutions of (5.31). The solutions G_η^G and G_η^w are increasing and convex in x and satisfy $\lim_{x \rightarrow \infty} \frac{G_\eta^c(x)}{x} = k_c^*$. Furthermore, see Figure 5.9, G_η^G is increasing in η . This is because, certainly when x is large, it is advantageous to stop, and the stopping rate increases as η increases. The picture for G^w as a function of η is very similar.

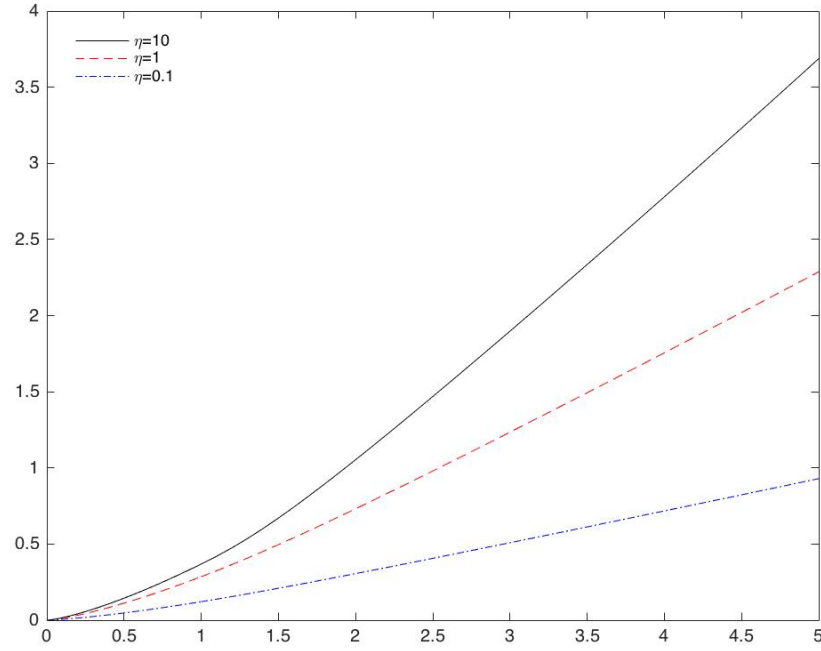


Figure 5.9: $(\beta, \mu, \sigma) = (5, 3, 2)$. G_η^G as a function of η . We see that G_η^G is increasing in η . We find a similar picture for G_η^w .

In Figure 5.10 we compare G_η^G with G_η^w . When $\eta = 0.1$ or $\eta = 1.0$ we find $G^G(x) > G^w(x)$ for all values of x . However, when $\eta = 10$ there is no universal relationship between G_η^G and G_η^w . We still find that $G_\eta^G(x) > G_\eta^w(x)$ for large x , but for small x the inequality is reversed. As we have found elsewhere, the feedback element implicit in the definition of G^G means that an increased value function increases the stopping rate, which can lower the value function in the region where g is small and stopping is not beneficial.

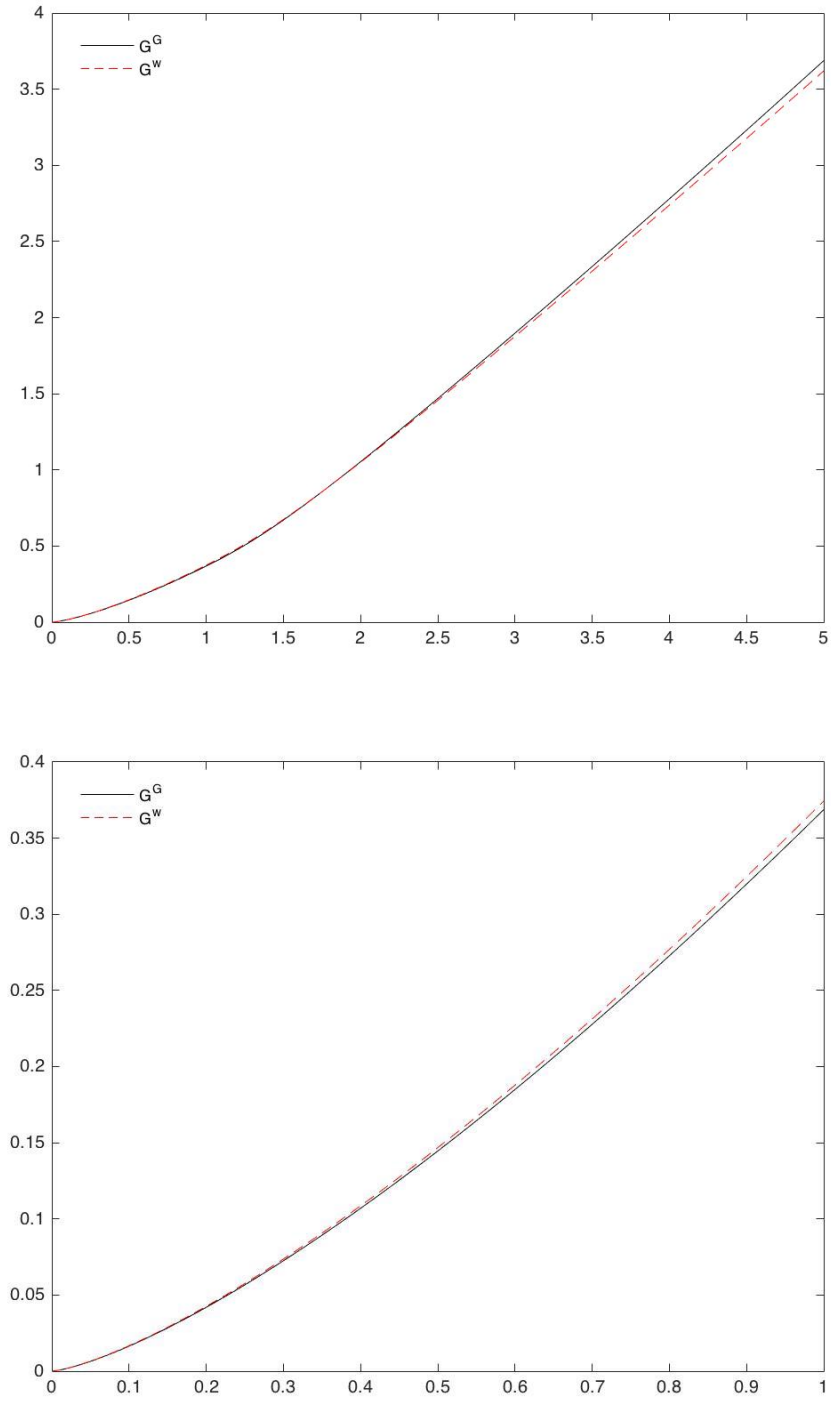


Figure 5.10: $(\beta, \mu, \sigma) = (5, 3, 2)$. A comparison of the value functions G_η^G and G_η^w when $\eta = 10$. For large x $G_\eta^G > G_\eta^w$; for small x (see the second graph), we find that $G_\eta^G < G_\eta^w$.

5.7 Conclusion

In the paradigm of randomising stopping rule, stochastic behaviours are derived from the model setup that stopping does not happen with probability one. The agents may not be sophisticated enough to compute the continuation value, and thus, are unable to make an unequivocal choice between stopping and continuing. Therefore, stopping probability at each decision point is based a stopping rule that takes into account the current stopping value and perceived continuation. We innovate the idea that the agent is aware of his future stochastic behaviours and is capable of computing the continuation value accordingly. This definition introduces non-linear feedback into the value function. Moreover, stopping can only happen at the event time of an exogenous Poisson process. This assumption is essential to avoid the degenerate case where stopping occurs immediately if stopping is allowed at any instant. However, we are able to analyse the limit case where the event of the Poisson process comes infinitely fast (or equivalently, stopping opportunity is available at any instant) by using a biased randomising stopping rule. And we have alternative interpretation for the agents' behaviours in the limit case: the agents will stop at the next event time of a Poisson process whose instant rate is dependent of the relative value of the current stopping and continuation value. In contrast to our dynamic CSC model where randomisation is not an assumption, the stochastic behaviour in this stopping problem is engineered in the setup of randomising stopping rule.

Chapter 6

Constrained Optimal Stopping, Liquidity and Effort

6.1 Introduction

Implicit in the classical version of the stopping problem defined in (1.1) and discussed in Chapter 2 is the idea that the agent can sell the asset (decide to invest, exercise the option) at any moment of their choosing, and for financial assets traded on an exchange this is a reasonable assumption. However, for other classes of assets, including those described as ‘real assets’ by, for example, Dixit and Pindyck (1994), this assumption may be less plausible. Therefore, in this chapter, we assume that the agent can only complete the sale if they can find a buyer, and candidate buyers are only available at certain isolated instants of time.

In this work we model the arrival of candidate purchasers as the event times of a Poisson process. When a candidate purchaser arrives the agent can choose to sell to that purchaser, or not; if a sale occurs then the problem terminates, otherwise the candidate purchaser is lost, and the problem continues. If the Poisson process has a constant rate, then the analysis falls into the framework studied by Dupuis and Wang (2005) and Lempa (2007).

Dupuis and Wang (2005) and Lempa (2007) discuss optimal stopping problems, but closely related is the work of Rogers and Zane (2000) in the context of portfolio optimisation. Rogers and Zane consider an optimal investment portfolio problem under the hypothesis that the portfolio can only be rebalanced at event times of a Poisson process of constant rate, see also Pham and Tankov (2008) and Ang, Papanikolaou and Westerfield (2014). The study of optimal stopping problems when the stopping times are constrained to be event times of an exogenous process is

relatively unexplored, but Guo and Liu (2005) study a problem in which the aim is to maximise a payoff contingent upon the maximum of an exponential Brownian motion and Menaldi and Robin (2016) extend the analysis of Dupuis and Wang (2005) to consider non-exponential inter-arrival times. As a generalisation of optimal stopping, Liang and Wei (2016) consider an optimal switching problem when the switching times are constrained to be event times of a Poisson process.

In this article we consider a more sophisticated model of optimal stopping under constraints in which the agent may expend effort in order to increase the frequency of the arrival times of candidate buyers. (Note that the problem remains an optimal stopping problem, since at each candidate sale opportunity the agent optimises between continuing and selling.) In our model the agent's instantaneous effort rate E_t affects the instantaneous rate Λ_t of the Poisson process, so that the candidate sale opportunities become the event times of an inhomogeneous Poisson process, where the agent chooses the rate. However, this effort is costly, and the agent incurs a cost per unit time which depends on the instantaneous effort rate. The objective of the agent is to maximise the expected discounted payoff net of the expected discounted costs. In particular, if $X = (X_t)_{t \geq 0}$ with $X_0 = x$ is the asset price process, g is the payoff function, β is the discount factor, $E = (E_t)_{t \geq 0}$ is the chosen effort process, $\Lambda = (\Lambda_t)_{t \geq 0}$ given by $\Lambda_t = \Psi(E_t)$ is the instantaneous rate of the Poisson process, C_E is the cost function so that the cost incurred per unit time is $C_E(E_t)$, and T_Λ is the set of event times of a Poisson process, rate Λ , then the objective of the agent is to maximise the objective function

$$\mathbb{E}^x \left[e^{-\beta\tau} g(X_\tau) - \int_0^\tau e^{-\beta s} C_E(E_s) ds \right] \quad (6.1)$$

over admissible effort processes E and T_Λ -valued stopping times τ . Our goal is to solve for the value function, the optimal stopping time and the optimal effort, as represented by the optimal control process E . In fact, typically it is possible to use the rate of the Poisson process as the control variable by setting $C(\Lambda_t) = C_E(E_t) = C_E \circ \Psi^{-1}(\Lambda_t)$. In the context of the problem it is natural to assume that Ψ and C_E are increasing functions, so that Ψ^{-1} exists, and C is increasing.

Our focus is on the case where X is an exponential Brownian motion, but the general case of a regular, time-homogeneous diffusion can be reduced to this case at the expense of slightly more complicated technical conditions. See Lempa [34] for a discussion in the constant arrival rate case. We begin by rigorously stating the form of the problem we will study. Then we proceed to solve for the effort process and stopping rule in (6.1). It turns out that there are two distinctive cases depending

on the shape of C or more precisely on the finiteness or otherwise of $\lim_{\lambda \uparrow \infty} \frac{C(\lambda)}{\lambda}$. Note that it is not clear *a priori* what shape $C = C_E \circ \Psi^{-1}$ should take, beyond the fact that it is increasing. Generally one might expect an increasing marginal cost of effort (convex C_E) and a law of diminishing returns to effort (concave Ψ) which would correspond to a convex C . But a partial reverse is also conceivable: effort expended below a threshold has little impact, and it is only once effort has reached a critical threshold that extra effort readily yields further stopping opportunities; in this case Ψ would be convex and C might be concave.

One outcome of our analysis is that the agent exerts effort to create a positive stopping rate only if they are in the region where stopping is optimal. Outside this region, they typically exert no effort, and there are no stopping opportunities. Typically therefore, (although we give a counterexample in an untypical case) the agent stops at the first occasion where stopping is possible and the optimal stopping element of the problem is trivial.

6.2 The set-up

We work on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = (\mathcal{F}_t)_{t \geq 0})$ which satisfies the usual conditions and which supports a Brownian motion and an independent Poisson process. On this space there is a regular, time-homogeneous diffusion process $X = (X_t)_{t \geq 0}$ driven by the Brownian motion. We will assume that X is exponential Brownian motion with volatility σ and drift μ and has initial value x ; then

$$\frac{dX_t}{X_t} = \sigma dW_t + \mu dt, \quad X_0 = x.$$

Here μ and σ are constants with $\mu < \beta$. The agent has a perpetual option with increasing payoff $g : \mathbb{R}_+ \mapsto \mathbb{R}_+$ of linear growth. In our examples g is an American call: $g(x) = (x - K)_+$. Then, in the classical setting, the problem of the agent would be to maximise $\mathbb{E}[e^{-\beta\tau} g(Y_\tau)]$ over stopping times τ . Note that the linear growth condition, together with $\mu < \beta$, is sufficient to ensure that this classical problem is well-posed.

We want to introduce finite liquidity into this problem, in the sense that we want to incorporate the phenomena that in order to sell the agent needs to find a buyer, and such buyers are in limited supply. In the simplest case buyers might arrive at event times of a time-homogeneous Poisson process with rate λ , and then at each event time of the Poisson process the agent faces a choice of whether to sell to this buyer at this moment or not; if yes then the sale occurs and the optimal stopping

problem terminates, if no then the buyer is irreversibly lost, and the optimal stopping problem continues. We want to augment this problem to allow the agent to expend effort (via networking, research or advertising) in order to increase the flow of buyers. There is a cost of searching in this way — the higher the effort the higher the rate of candidate stopping times but also the higher the search costs. Note that once the asset is sold, effort expended on searching ceases, and search costs thereafter are zero by fiat.

Let \mathcal{A}_E be the set of admissible effort processes. We assume that $E \in \mathcal{A}_E$ if $E = (E_t)_{t \geq 0}$ is an adapted process such that $E_t \in I_E$ for all $t \in [0, \infty)$ where $I_E \subset \mathbb{R}_+$ is an interval which is independent of time. Then, since $\Lambda_t = \Psi(E_t)$ we find $E \in \mathcal{A}_E$ if and only if $\Lambda \in \mathcal{A}$ where $\Lambda \in \mathcal{A}$ if Λ is adapted and $\Lambda_t \in I$ for all t where $I = \Psi(I_E)$. Note that I is an interval in \mathbb{R}_+ , and we take the lower and upper endpoints to be $\underline{\lambda}$ and $\bar{\lambda}$ respectively.

Recall that T_Λ is the set of event times of an inhomogeneous Poisson process with rate Λ . Then $T^\Lambda = \{T_1^\Lambda, T_2^\Lambda, \dots\}$ where $0 < T_1^\Lambda$ and $T_n^\Lambda < T_{n+1}^\Lambda$ almost surely. Let $\mathcal{T}(T_\Lambda)$ be the set of T_Λ -valued stopping times and let \mathcal{A} be the set of admissible rate functions. Then, after a change of independent variable the problem is to find

$$H(x) = \sup_{\Lambda \in \mathcal{A}} \sup_{\tau \in \mathcal{T}(T_\Lambda)} \mathbb{E}^x \left[e^{-\beta\tau} g(X_\tau) - \int_0^\tau e^{-\beta s} C(\Lambda_s) ds \right], \quad (6.2)$$

together with the optimal rate function $\Lambda^* = (\Lambda_t^*)_{t \geq 0}$ and optimal stopping rule $\tau^* \in \mathcal{T}(T_\Lambda)$.

In addition to the set of admissible controls, we also consider the subset of integrable controls $\mathcal{I} \subseteq \mathcal{A}$ where $\Lambda \in \mathcal{I} = \mathcal{I}(I, C)$ is an adapted process with $\Lambda_t \in I$ for which $\mathbb{E}^x[\int_0^\infty e^{-\beta s} C(\Lambda_s) ds] < \infty$. As mentioned above we have that $\mathbb{E}^x[e^{-\beta\tau} g(X_\tau)] < \infty$ for any admissible Λ and any stopping rule, and hence there is no loss of generality in restricting the search for the optimal rate function to the set of integrable controls.

The stopping rule is easily identified in feedback form. Let $T_\Lambda^0 = T_\Lambda \cup \{0\}$ and let H^0 be the value of the problem *conditional on there being a buyer available at time 0*, so that

$$H^0(x) = \sup_{\Lambda \in \mathcal{A}} \sup_{\tau \in \mathcal{T}(T_\Lambda^0)} \mathbb{E}^x \left[e^{-\beta\tau} g(X_\tau) - \int_0^\tau e^{-\beta s} C(\Lambda_s) ds \right].$$

Then, it is optimal to stop immediately if and only if the value of stopping is at

least as large as the value of continuing and

$$H^0(x) = \max\{g(x), H(x)\}.$$

It follows that if $\Lambda = (\Lambda_t)_{t \geq 0}$ is a fixed admissible rate process, and if H_Λ^0 and H_Λ denote the respective value functions then, writing $T_1 = T_1^\Lambda$ for the first event time of the Poisson process with rate Λ ,

$$\begin{aligned} H_\Lambda(x) &= \sup_{\tau \in \mathcal{T}(T_\Lambda)} \mathbb{E}^x \left[e^{-\beta\tau} g(X_\tau) - \int_0^\tau e^{-\beta s} C(\Lambda_s) ds \right] \\ &= \sup_{\tau \in \mathcal{T}(T_\Lambda)} \mathbb{E}^x \left[e^{-\beta T_1} \mathbb{E} \left[e^{-\beta(\tau - T_1)} g(X_\tau) - \int_{T_1}^\tau e^{-\beta(s - T_1)} C(\Lambda_s) ds \middle| \mathcal{F}_{T_1} \right] - \int_0^{T_1} e^{-\beta s} C(\Lambda_s) ds \right] \\ &= \mathbb{E}^x \left[e^{-\beta T_1} H_\Lambda^0(X_{T_1}) - \int_0^\infty I_{\{s < T_1\}} e^{-\beta s} C(\Lambda_s) ds \right] \\ &= \mathbb{E}^x \left[\int_0^\infty \Lambda_s e^{-\int_0^s \Lambda_u du} e^{-\beta s} H_\Lambda^0(X_s) ds - \int_0^\infty e^{-\int_0^s \Lambda_u du} e^{-\beta s} C(\Lambda_s) ds \right] \\ &= \mathbb{E}^x \left[\int_0^\infty e^{-(\beta s + \int_0^s \Lambda_u du)} (\Lambda_s H_\Lambda^0(X_s) - C(\Lambda_s)) ds \right]. \end{aligned}$$

Taking a supremum over admissible rate processes $\Lambda \in \mathcal{A}$ we find

$$H(x) = \sup_{\Lambda \in \mathcal{A}} \mathbb{E}^x \left[\int_0^\infty e^{-\int_0^t (\beta + \Lambda_s) ds} (\Lambda_t H_\Lambda^0(X_t) - C(\Lambda_t)) dt \right],$$

and this is the problem we aim to solve. Writing Λ^* for the optimal rate process we expect H to solve

$$H(x) = \mathbb{E}^x \left[\int_0^\infty e^{-\int_0^t (\beta + \Lambda_s^*) ds} (\Lambda_t^* \{g(X_t) \vee H(X_t)\} - C(\Lambda_t^*)) dt \right].$$

Note that $H_\Lambda^0 \leq H^0$ and the equality is attained by Λ^* .

6.2.1 Some results for classical problems

For future reference we record some results for classical problems in which agents can stop at any instant.

First, let $\mathcal{T}([0, \infty))$ be the set of all stopping times and define

$$w^K(x) := \sup_{\tau \in \mathcal{T}([0, \infty))} \mathbb{E}^x [e^{-\beta\tau} (X_\tau - K)_+].$$

(Imagine a standard, perpetual, American-style call option with strike K , though valuation is not taking place under the equivalent martingale measure.) Classical arguments (McKean (1965), Peskir and Shiryaev (2006)) give that $0 < w^K < x$ (the upper bound holds since we are assuming $\beta > \mu$) and that there exists a constant

$L = \frac{\theta}{\theta-1}K$ where $\theta = (\frac{1}{2} - \frac{\mu}{\sigma^2}) + \sqrt{(\frac{1}{2} - \frac{\mu}{\sigma^2})^2 + \frac{2\beta}{\sigma^2}}$ such that

$$w^K(x) = \begin{cases} (x - K)_+, & x > L; \\ (L - K)L^{-\theta}x^\theta, & 0 < x \leq L. \end{cases}$$

For future reference set $\phi = (\frac{1}{2} - \frac{\mu}{\sigma^2}) - \sqrt{(\frac{1}{2} - \frac{\mu}{\sigma^2})^2 + \frac{2\beta}{\sigma^2}}$. Then $\phi < 0 < 1 < \theta$ and θ and ϕ are the roots of $Q_0 = 0$ where $Q_\lambda(\psi) = \frac{1}{2}\sigma^2\psi(\psi-1) + \mu\psi - (\beta + \lambda)$.

Second, define

$$w_{K,\epsilon,\delta}(x) = \sup_{\tau \in \mathcal{T}([0,\infty))} \mathbb{E}^x \left[e^{-\beta\tau} \{(X_\tau - K)_+ - \epsilon\} - \delta \int_0^\tau e^{-\beta s} ds \right]. \quad (6.3)$$

(Imagine a perpetual, American-style call option with strike K , in which the agent pays a fee or transaction cost ϵ to exercise the option, and pays a running cost δ per unit time until the option is exercised.) Note that $w_{K,0,0} \equiv w^K$. It turns out that there are two cases.

In the first case of $\epsilon \geq \delta/\beta$, when X is small it is more cost effective to pay the running cost indefinitely than to pay the exercise fee. We find

$$w_{K,\epsilon,\delta}(x) = w^{K+\epsilon-\delta/\beta}(x) - \delta/\beta.$$

We obtain the continuation region $\mathcal{C} = (0, L_{\epsilon,\delta})$ and stopping region $\mathcal{S} = [L_{\epsilon,\delta}, \infty)$ where $L_{\epsilon,\delta} = \frac{\theta}{\theta-1}(K + \epsilon - \frac{\delta}{\beta})$.

In the second case of $\epsilon < \delta/\beta$, when X is small and the option is out-of-the-money, it is cost-effective to stop immediately, even though the payoff is zero, because paying the fee is cheaper than paying the running cost indefinitely. Thus, we have

$$v(x) = -\epsilon, \quad x \in (0, l^*)$$

for some $l^* = l_{\epsilon,\delta}^* < K$ to be determined. On the other hand, when X is large and the option is in-the-money, the payoff of stopping is again larger than the value of continuing. Thus, we have

$$v(x) = x - K - \epsilon, \quad x \in (L^*, \infty)$$

for some $L^* = L_{\epsilon,\delta}^* > K$ to be determined. Thus the stopping region is $\mathcal{S} = (0, l^*) \cup (L^*, \infty)$. And the continuation region is $\mathcal{C} = (l^*, L^*)$ and we have,

$$\mathcal{L}v(x) - \beta v(x) = \delta, \quad x \in (l^*, L^*).$$

The value of l^* and L^* can be determined by smooth-fit principle. Define stopping time $\tau^* = \inf\{t > 0; X_t \notin (l^*, L^*)\}$. The proof of the optimality of τ^* and $v = w_{K,\epsilon,\delta}$ is similar to the proof in Theorem 1.

Returning to our problem with limited stopping opportunities, one immediate observation is that $H(x) \leq w_K(x)$. Conversely, if $\Lambda \equiv 0$ is admissible then $H(x) \geq -\frac{C(0)}{\beta}$.

6.3 Cost functions

In this section, we list several different types of cost functions that we will discuss in the following sections.

In Section 6.4.1 and Section 6.4.3, we present the optimal control when the cost function is quadratic, that is

$$C(\lambda) = a + b\lambda + c\frac{\lambda^2}{2}$$

with $a \geq 0$, $b \geq 0$ and $c > 0$. And we include numerical analysis for the special case where $a = 0 = b$.

In Section 6.5, we prove the optimal control when the cost function is concave. And we include numerical analysis for the case where

$$C(\lambda) = \sqrt{\lambda}$$

In Section 6.6, we provide further numerical analysis on stopping behaviours under various cost functions. More specifically, we include

- $C_0(\lambda) = c\frac{\lambda^2}{2}$
- $C_b(\lambda) = c\frac{\lambda^2}{2} + b\lambda$
- $C_1(\lambda) = c\frac{\lambda^2}{2} + a$
- $C_>(\lambda) = c\frac{\lambda^2}{2} + a\mathbb{I}_{\{\lambda > 0\}}$

6.4 Quadratic cost functions

6.4.1 Heuristics

From the Markovian structure of the problem we expect that the (unknown) value function H and optimal rate function Λ^* are time-homogeneous functions of the asset price only.

Let $M^\Lambda = (M_t^\Lambda)_{t \geq 0}$ be given by

$$M_t^\Lambda = e^{-\int_0^t (\beta + \Lambda_s) ds} H(X_t) + \int_0^t e^{-\int_0^u (\beta + \Lambda_s) ds} [\Lambda_u H^0(X_u) - C(\Lambda_u)] du,$$

and let \mathcal{L}^X denote the generator of X so that $\mathcal{L}^X f = \frac{\sigma^2}{2} f'' + \mu f'$. Assume that the value function under the optimal strategy H is C^2 . Then, by Itô's formula,

$$dM_t^\Lambda = e^{-\int_0^t (\beta + \Lambda_s) ds} \left\{ \left(\mathcal{L}^X H(X_t) - (\beta + \Lambda_t) H(X_t) + \Lambda_t (H^0(X_t) - C(\Lambda_t)) \right) dt + \sigma X_t H'(X_t) dW_t \right\}.$$

We expect that M^Λ is a super-martingale for any choice of Λ , and a martingale for the optimal choice. Thus we expect

$$\mathcal{L}^X H(X_t) - \beta H(X_t) - \inf_{\Lambda_t} \{C(\Lambda_t) - \Lambda_t [H^0(X_t) - H(X_t)]\} = 0.$$

Let $\tilde{C} : \mathbb{R}_+ \mapsto \mathbb{R}$ be the concave conjugate of C so that $\tilde{C}(z) = \inf_{\lambda \geq 0} \{C(\lambda) - \lambda z\}$. Then we find that H solves

$$\mathcal{L}^X H - \beta H - \tilde{C}(H^0 - H) = 0, \quad (6.4)$$

and a best choice of rate function is $\Lambda_t^* = \Lambda^*(X_t)$ where

$$\Lambda^*(x) = \Theta(H^0(x) - H(x)) \quad (6.5)$$

and $\Theta(z) = \arg\inf_{\lambda} \{C(\lambda) - \lambda z\}$. Note that $H^0 - H = (g - H)_+$ and that (6.4) is a second order differential equation and will have multiple solutions. The boundary behaviour near zero and infinity will determine which solution fits the optimal stopping problem.

6.4.2 First Example

Suppose $g(x) = (x - K)_+$ for fixed $K > 0$. Using terminology from the study of American options and optimal stopping we say that if $X_t > K$ then the process is in-the-money, if $X_t < K$ then the process is out-of-the-money and the region in the domain of X where $\Lambda^*(X)$ is zero is the continuation region \mathcal{C} , and $\mathcal{S} := \mathbb{R}^+ \setminus \mathcal{C}$ is the selling region.

Suppose $\underline{\lambda} = 0$ and $\bar{\lambda} = \infty$, then the range of possible values for the rate process is $I = [0, \infty]$ and consider a quadratic cost function $C(\lambda) = a + b\lambda + c\frac{\lambda^2}{2}$ with $a \geq 0$, $b \geq 0$ and $c > 0$. Then, we have

$$\tilde{C}(z) = a - \frac{(z - b)^2}{2c}. \quad (6.6)$$

It follows that,

$$\mathcal{L}^X h - \beta h = a - \frac{[(g - h)_+ - b]^2}{2c}, \quad (6.7)$$

Consider first the behaviour of the value function near zero. If $a = 0$ then $C(0) = 0$, and when X is close to zero the agent may choose not to search for buyers, a strategy which incurs zero cost. There is little chance of the process ever being in-the-money, but nonetheless the agent delays sale indefinitely. We expect that the continuation region is $(0, L^*)$ for some threshold L^* .

Now suppose $a > 0$. Now there is a cost to delaying the sale, even when $\Lambda = 0$. If X is small then it is preferable to sell the asset even though the process is out-of-the-money, because in our problem there are no search costs once the asset is sold. In this case we expect the agent to search for buyers when X is small, in order to reduce further costs. Then the continuation region will be (ℓ^*, L^*) for some $0 < \ell^* < K < L^* < \infty$.

Consider now the behaviour for large x . In this case we can look for an expansion for the solution of (6.4) of the form

$$H(x) = A_1 x + A_{1/2} \sqrt{x} + A_0 + O(x^{-1/2}) \quad (6.8)$$

for constants A_1 , $A_{1/2}$ and A_0 to be determined. Using the fact that $H(x) \leq w_K(x)$ so that H is of at most linear growth. The reason we expect there is a square-root term of x is due to the fact that $A_{1/2} \sqrt{x}$ becomes a linear term of x after taking the square on the right-hand-side of (6.7). Moreover, we expect $A_1 = 1$ so that the quadratic terms of x cancels on the right-hand-side of (6.7). By plugging (6.8) into (6.7) and setting the coefficients of x and \sqrt{x} terms to be 0, we obtain,

$$H(x) = x - \sqrt{2c(\beta - \mu)} \sqrt{x} - \left\{ K + b - c \left[\beta - \frac{\mu}{2} + \frac{\sigma^2}{8} \right] \right\} + \dots \quad (6.9)$$

Numerical results (see Figure 6.1) show that this expansion is very accurate for large x .

Purely quadratic cost: $a = 0 = b$

In this case we expect that the continuation region is $(0, L^*)$ for a threshold level L^* to be determined. For a general threshold L , and writing H_L for the solution to (6.4) with $H(0) = 0$ and $H(L) = L - K$ we find that H_L solves

$$\mathcal{L}^X h - \beta h = \frac{1}{2c} (\{g - h\}_+)^2, \quad (6.10)$$

and then that $H_L(x) = \frac{L-K}{L^\theta} x^\theta$ on $x \leq L$. On (L, ∞) , H_L solves (6.10) subject to $H_L(L) = (L - K)$ and $H'_L(L) = \theta \frac{L-K}{L}$. This procedure gives us a family $(H_L)_{L \geq K}$ of potential value functions, each of which is C^1 . Finally we can determine the threshold level L by choosing the value L^* for which H_{L^*} has linear growth at infinity.

The linear growth solution H_{L^*} is shown in Figure 6.1, both for large x and for moderate x . From We also see that the expansion for H given in (6.9) gives a good approximation of our numerical solution for large x .

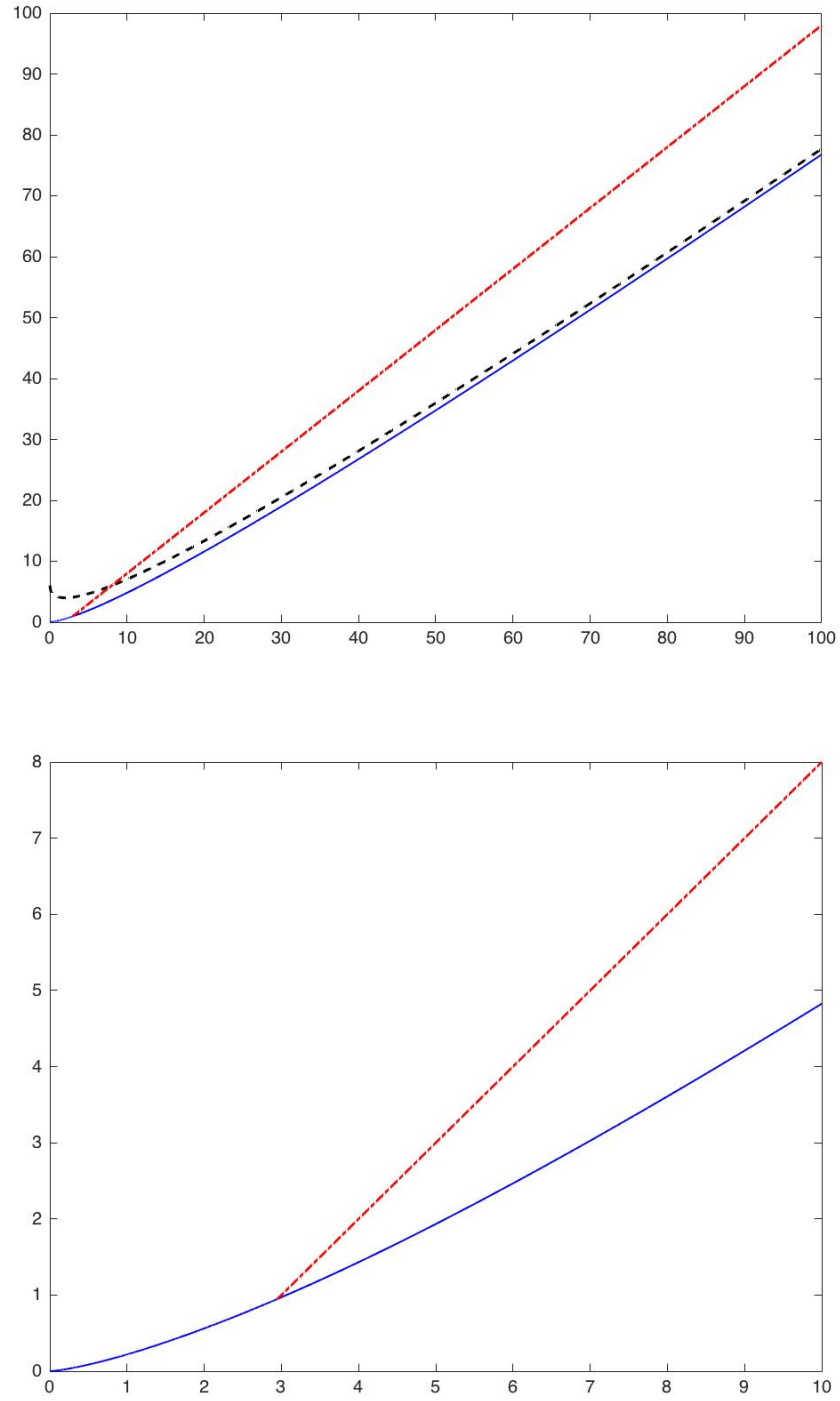


Figure 6.1: $(\beta, \mu, \sigma, K, a, b, c) = (5, 3, 2, 1, 0, 0, 2)$. In both sub-figures the solid curved line represents H_{L^*} ; the straight line represents $g \vee H_{L^*}$ on $\{x : g(x) \geq H_{L^*}(x)\}$ and the dashed line in the top sub-figure is the expansion for H in (6.9).

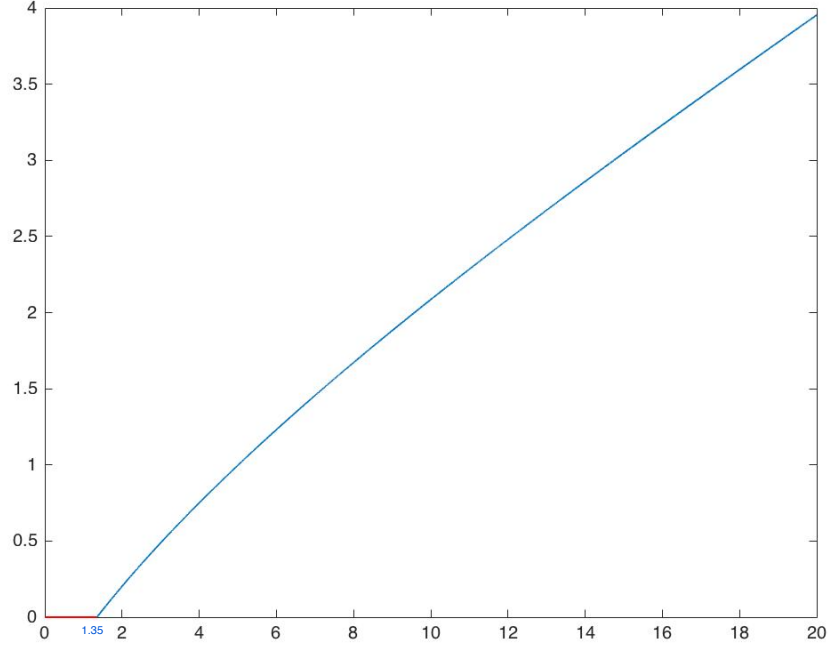


Figure 6.2: $(\beta, \mu, \sigma, K, a, b, c) = (5, 3, 2, 1, 0, 0, 2)$; this figure plots the optimal control Λ^* given by (6.5) as a function of wealth level x . The optimal threshold is seen to be at $L^* = 1.35$.

Figure 6.2 we see that the continuation region is $\mathcal{C} = (0, 1.35)$ and that the stopping region $\mathcal{S} = [1.35, \infty)$. We see that Λ^* is zero on the continuation region \mathcal{C} and that Λ^* is increasing and concave on the stopping region \mathcal{S} . The agent behaves rationally in the sense that on the continuation region where continuing is worth more than stopping, the agent is unwilling to stop and this is reflected by the minimal efforts spent on searching (i.e. $\Lambda^*(x) = 0, \forall x \in \mathcal{C}$); similarly, on the stopping region, stopping is getting more and more valuable relative to continuing as the price process gets deeper in-the-money, and the agent is incentivised to spend more effort on searching for stopping opportunities.

The analysis on varying parameters is presented as follow.

Cost coefficient c

We can see that from Figure 6.3 that both the value function H_{L^*} and the optimal stopping threshold L^* are getting smaller as the cost coefficient increases. Figure 6.4 shows that for small x , the agent spends more efforts on searching for trading opportunities when the cost coefficient is larger. This phenomenon results from the

fact that agents with higher cost coefficients have a smaller stopping threshold; on the other hand, for large x , the agent spends less efforts on searching trading opportunities (choosing a smaller rate for the Poisson process) when the cost coefficient is larger. This is due to the fact that larger cost coefficient makes it more costly for searching, and hence the larger cost makes the optimal control smaller.

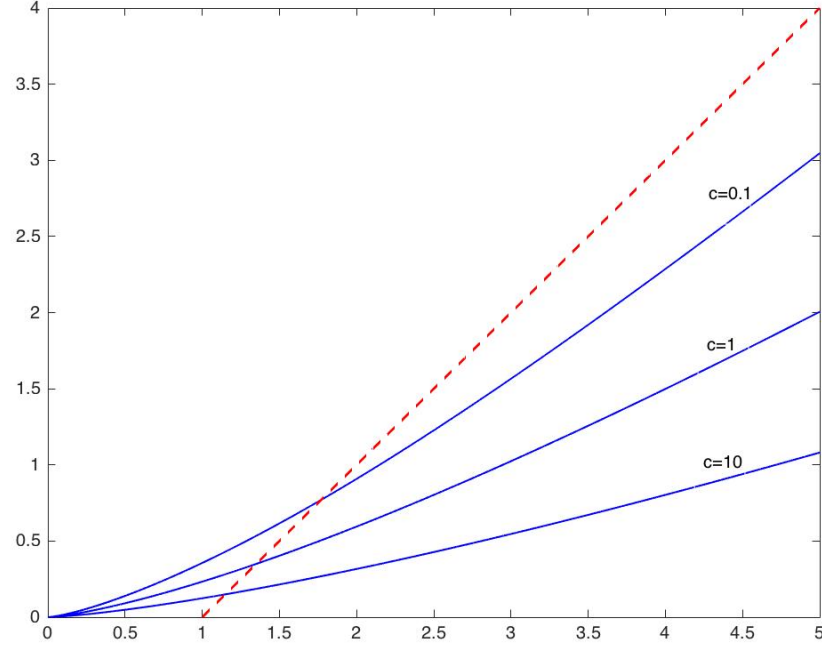


Figure 6.3: These solid lines depicts the value function H_L^* under different cost coefficient c ; the dashed line represents $g(x)$; $(\beta, \mu, \sigma, K) = (5, 3, 2, 1)$; the optimal stopping threshold is indicated by the crosspoint of each solid line with the dashed line.

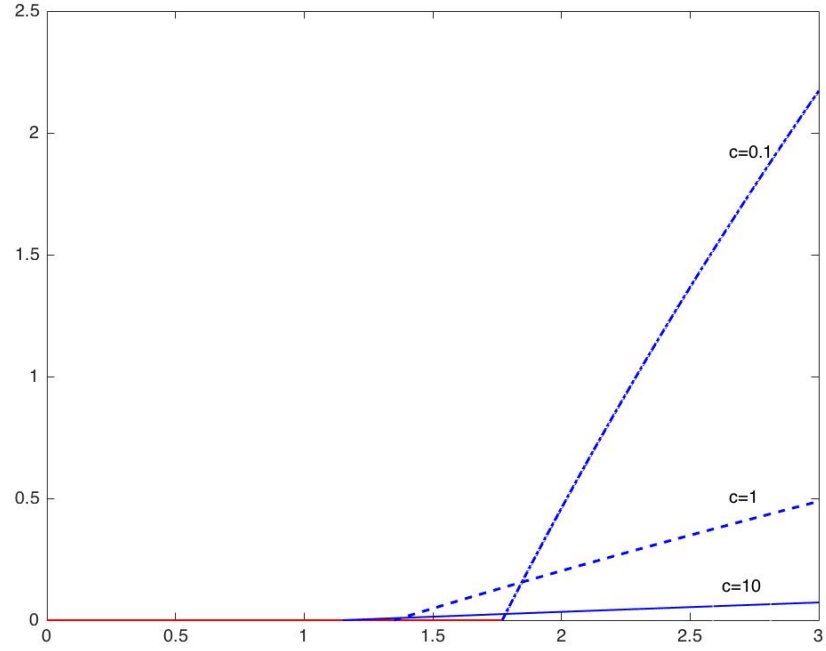


Figure 6.4: The optimal controls Λ^* under different cost coefficients c .

Discount factor β

We can see from Figure 6.5 that the value function H_{L^*} and the optimal stopping threshold L_* are getting smaller as the discounting coefficient increases. Figure 6.6 shows that the agent spends more efforts on searching (choosing larger rate for Poisson process) given a larger discounting factor. This phenomenon is due to the fact that continuing becomes less valuable as β increases. The agent has incentive to spend more efforts on searching opportunities to stop for larger β .

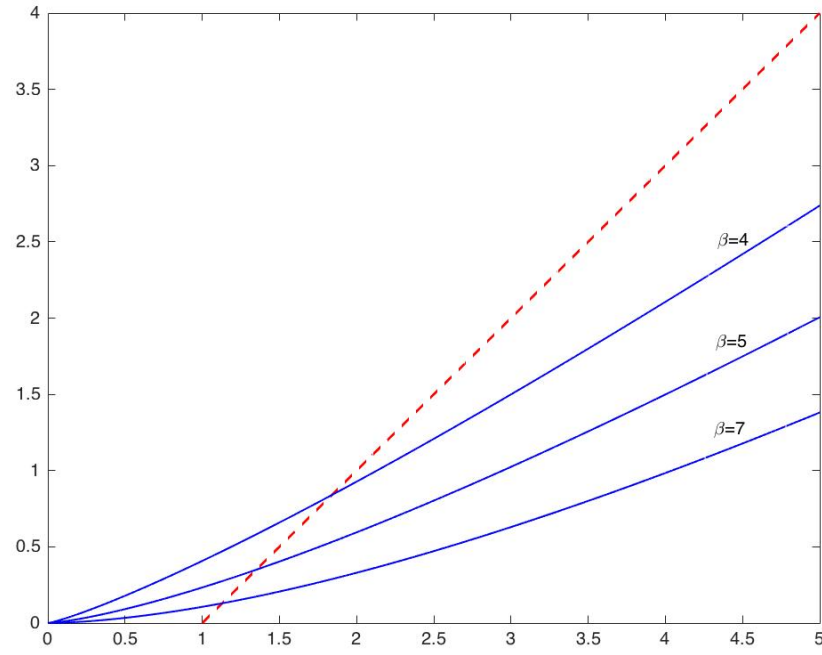


Figure 6.5: these blue lines depicts the value function H_L^* for different discount parameters $\beta = 4, 5, 7$ from top to bottom; the red dashed line represents $g(x); (\mu, \sigma, K, c) = (3, 2, 1, 1)$; the optimal stopping threshold is indicated by the crosspoint of each blue line with the red dashedline in (a)

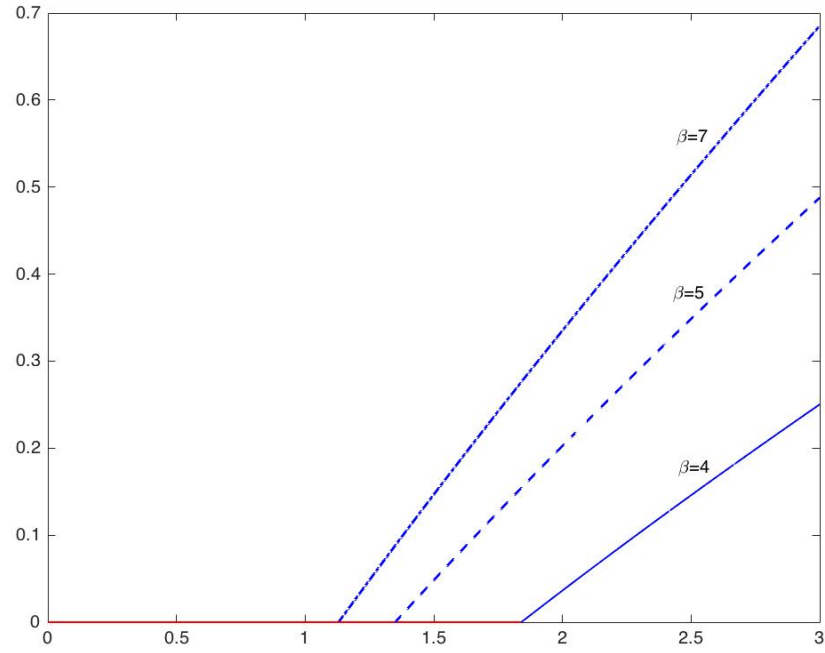


Figure 6.6: the optimal control Λ^* under different discounting factors.

Drift coefficient μ

We can see from Figure 6.7 that the value function H_{L^*} and the optimal stopping threshold L^* are getting smaller as drift coefficient decreases. Figure 6.8 shows that the agent spends more efforts on searching (choosing larger rate for Poisson process) given a smaller drift coefficient. Hence, this is consistent with our intuition that the impact of a drift coefficient leads to the opposite direction compared with that of a discount factor.

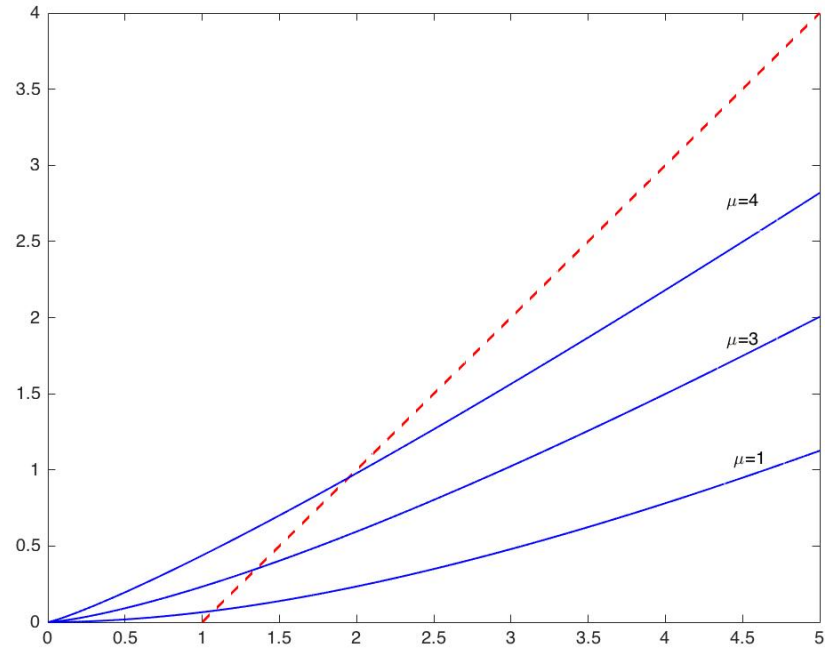


Figure 6.7: these solid lines depict the value function H_L^* for different drift coefficient $\mu = 4, 3, 1$ from top to bottom; the dashed line represents $g(x)$; $(\beta, \sigma, K, c) = (5, 2, 1, 1)$; the optimal stopping threshold is indicated by the crosspoint of each solid line with the dashed line.

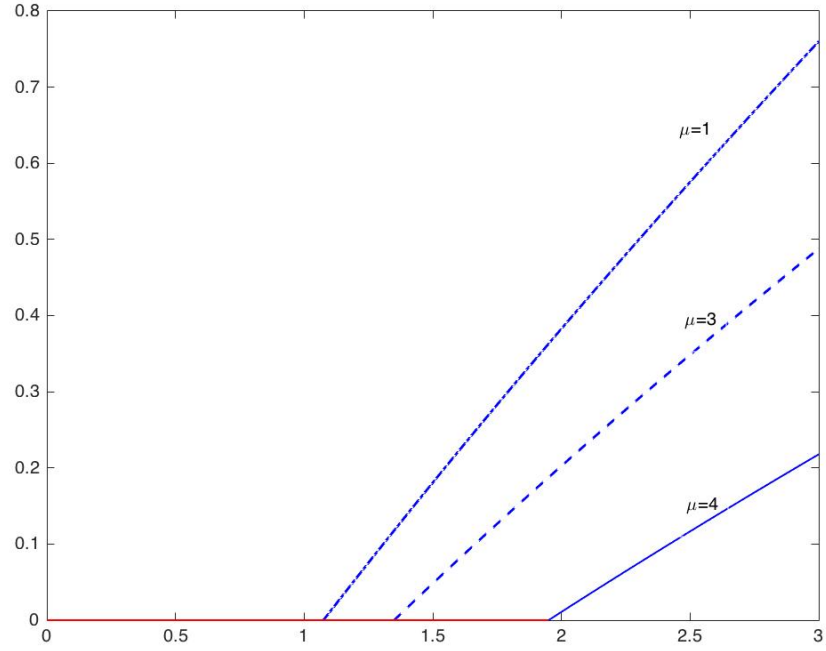


Figure 6.8: the optimal control Λ^* under different drift coefficients.

Volatility coefficient σ

We can see from Figure 6.9 that the value function H_L^* and the optimal stopping threshold L^* are getting larger as volatility increases. This is consistent with the intuition that larger volatility the underlying process has, the more valuable the option is, which results from the convexity of the option payoff; Figure 6.10 shows that the agent spends more efforts on searching (choosing larger rate for Poisson process) given a smaller volatility. This phenomenon is due to the fact smaller volatility implies continuing is less valuable. The agent has incentive to spend more efforts on searching opportunities to stop.

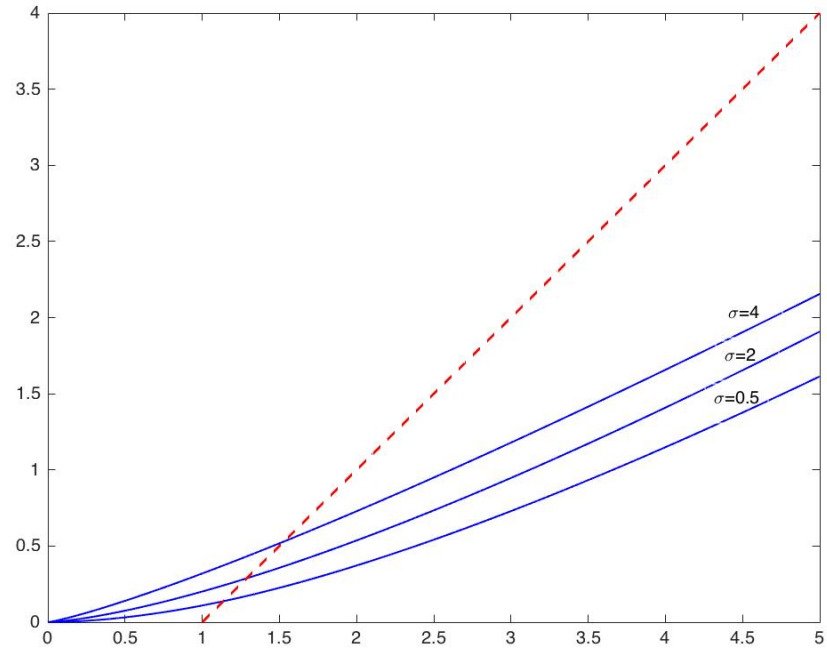


Figure 6.9: these solid lines depicts the value function H_L^* for different volatility $\sigma = 4, 2, 0.5$ from top to bottom; the dashed line represents $g(x)$; $(\beta, \mu, K, c) = (5, 3, 1, 1)$; the optimal stopping threshold is indicated by the crosspoint of each solid line with the dashed line.

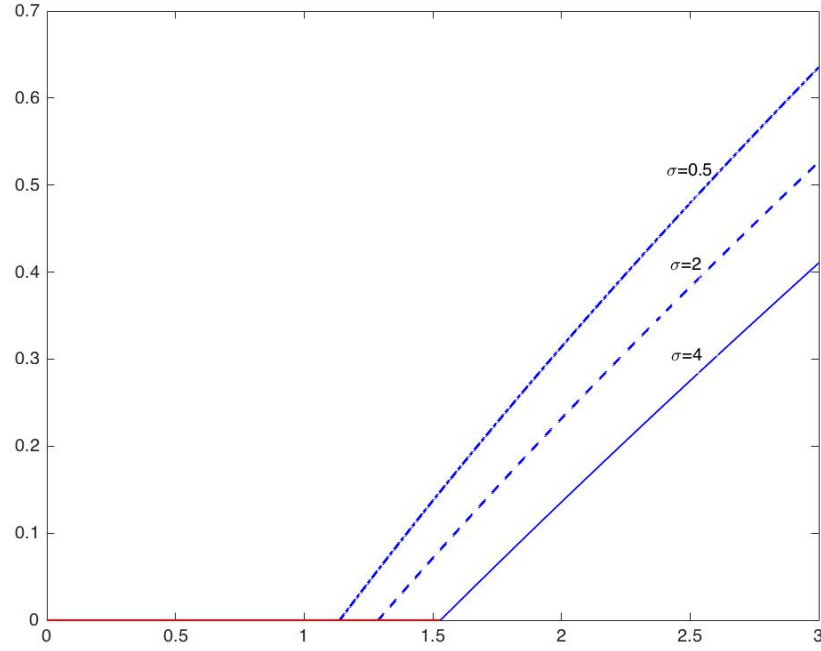


Figure 6.10: the optimal control Λ^* under different volatility coefficients.

We discuss the cases of $a > 0$ and $b > 0$ in Section 6.6.

6.4.3 Verification

In this section we show that the heuristics are correct, and that the value to the stochastic problem is given by the appropriate solution of the differential equation. Although the details are different, the structure of the proof follows Dupuis and Wang (2005).

Suppose, as throughout, that X is exponential Brownian motion with $\mu < \beta$ and g is of linear growth.

Definition 1. (τ, Λ) is admissible if Λ is a non-negative, I -valued, adapted process and $\tau \in \mathcal{T}(T^\Lambda)$.

Note that a consequence of the definition is that we insist that $\tau \leq T_\infty^\Lambda := \lim_n T_n^\Lambda$. Moreover, we may have $T_k = \infty$: in this case we may take $\tau = \infty$, whence we have $e^{-\beta\tau}g(X_\tau) = 0$ noting that $\lim_{t \uparrow \infty} e^{-\beta t}g(X_t) = 0$ almost surely.

Definition 2. (τ, Λ) is integrable if (τ, Λ) is admissible and $\mathbb{E}[\int_0^\tau e^{-\beta s}C(\Lambda_s)ds] < \infty$.

Clearly, if (τ, Λ) is integrable, then (T_1^Λ, Λ) is integrable.

Lemma 13. *Let G be an increasing, convex solution to*

$$\mathcal{L}^X G - \beta G - \tilde{C}((g - G)_+) = 0, \quad (6.11)$$

and suppose that G is of at most linear growth. Set $G^0 = G \vee g$.

Then for any integrable, admissible strategy (τ, Λ) ,

$$G(x) \geq \mathbb{E}^x \left[e^{-\beta T_1^\Lambda} G^0(X_{T_1^\Lambda}) I_{\{T_1^\Lambda < \infty\}} - \int_0^{T_1^\Lambda} e^{-\beta s} C(\Lambda_s) ds \right]. \quad (6.12)$$

Proof. Since g and G are of linear growth we may assume $G^0(x) \leq \kappa_0 + \kappa_1 x$ for some constants $\kappa_i \in (0, \infty)$.

Let $Z_t = e^{-\beta t - \int_0^t \Lambda_s ds} G(X_t) - \int_0^t e^{-\beta s - \int_0^s \Lambda_u du} F_s ds$ where

$$F_s = F(g(X_s), G(X_s), \Lambda_s) := (g(X_s) - G(X_s))_+ \Lambda_s + \tilde{C}((g(X_s) - G(X_s))_+) \leq C(\Lambda_s).$$

Then, using the definition of G

$$\begin{aligned} dZ_t &= e^{-\beta t - \int_0^t \Lambda_s ds} \left\{ -(\beta + \Lambda_t)G + \mathcal{L}^X G - (g - G)_+ \Lambda_t - \tilde{C}((g - G)_+) \right\} dt + dN_t \\ &= e^{-\beta t - \int_0^t \Lambda_s ds} \left\{ -\Lambda_t [G + (g - G)_+] \right\} dt + dN_t \\ &= -e^{-\beta t - \int_0^t \Lambda_s ds} \Lambda_t G^0(X_t) dt + dN_t \end{aligned}$$

where $N_t = \int_0^t e^{-\beta s - \int_0^s \Lambda_u du} \sigma X_s G'(X_s) dW_s$. Our hypotheses on G allow us to conclude that $N = (N_t)_{t \geq 0}$ is a martingale.

It follows that $Z_0 = \mathbb{E}[Z_t + \int_0^t e^{-\beta s - \int_0^s \Lambda_u du} \Lambda_s G^0(X_s) ds]$ or equivalently

$$\begin{aligned} G(x) &= \mathbb{E}^x \left[e^{-\beta t - \int_0^t \Lambda_s ds} G(X_t) + \int_0^t e^{-\beta s - \int_0^s \Lambda_u du} (\Lambda_s (g(X_s) \vee G(X_s)) - F_s) ds \right] \\ &\geq \mathbb{E}^x \left[e^{-\beta t - \int_0^t \Lambda_s ds} G(X_t) + \int_0^t e^{-\beta s - \int_0^s \Lambda_u du} (\Lambda_s G^0(X_s) - C(\Lambda_s)) ds \right]. \quad (6.13) \end{aligned}$$

Since X is geometric Brownian motion and $\beta > \mu$ we have that $X^{\beta,*} := \sup_{u \geq 0} \{e^{-\beta u} X_u\}$ is in L^1 . Then

$$\begin{aligned} e^{-\beta t - \int_0^t \Lambda_s ds} G(X_t) &\leq \kappa_0 + \kappa_1 X^{\beta,*}, \\ \int_0^t e^{-\beta s - \int_0^s \Lambda_u du} \Lambda_s G^0(X_s) ds &\leq (\kappa_0 + \kappa_1 X^{\beta,*}) \int_0^t \Lambda_s e^{-\int_0^s \Lambda_u du} ds \leq \kappa_0 + \kappa_1 X^{\beta,*}, \\ \int_0^t e^{-\beta s - \int_0^s \Lambda_u du} C(\Lambda_s) ds &\leq \int_0^\infty e^{-\beta s - \int_0^s \Lambda_u du} C(\Lambda_s) ds, \end{aligned}$$

and, since (T_1^Λ, Λ) is integrable by hypothesis,

$$\mathbb{E} \left[\int_0^\infty e^{-\beta s - \int_0^s \Lambda_u du} C(\Lambda_s) ds \right] = \mathbb{E} \left[\int_0^{T_1^\Lambda} e^{-\beta s} C(\Lambda_s) ds \right] < \infty.$$

Notice that

$$\lim_{t \uparrow \infty} \mathbb{E}^x \left[e^{-\beta t - \int_0^t \Lambda_s ds} G(X_t) \right] = 0.$$

Let $t \rightarrow \infty$ and by Dominated Convergence we have,

$$G(x) \geq \mathbb{E}^x \left[\int_0^\infty e^{-\beta s - \int_0^s \Lambda_u du} (\Lambda_s G^0(X_s) - C(\Lambda_s)) ds \right].$$

Then, by Fubini's theorem, we obtain (6.12). □

Lemma 14. *Let (τ, Λ) be an integrable strategy. Define $Y = (Y_n)_{n \geq 0}$ by*

$$Y_n = e^{-\beta(T_n^\Lambda \wedge \tau)} G^0(X_{T_n^\Lambda \wedge \tau}) I_{\{T_n^\Lambda \wedge \tau < \infty\}} - \int_0^{T_n^\Lambda \wedge \tau} e^{-\beta s} C(\Lambda_s) ds$$

where $T_0^\Lambda = 0$. Define $\mathcal{G}_n = \mathcal{F}_{T_n^\Lambda}$ and set $\mathbb{G} = (\mathcal{G}_n)_{n \geq 0}$.

Then Y is a uniformly integrable $(\mathcal{G}_n)_{n \geq 0}$ -supermartingale.

Proof. We have

$$|Y_n| \leq \kappa_0 + \kappa_1 X^{\beta,*} + \int_0^\tau e^{-\beta s} C(\Lambda_s) ds \in L^1.$$

Moreover, on $T_{n-1}^\Lambda < \infty$ and $\tau > T_{n-1}^\Lambda$, writing \tilde{T} as shorthand for $T_n^\Lambda - T_{n-1}^\Lambda$ and using $\tau \geq T_n^\Lambda$ and Lemma 13 for the crucial first inequality,

$$\begin{aligned} \mathbb{E}[Y_n | \mathcal{G}_{n-1}] &= e^{-\beta T_{n-1}^\Lambda} \mathbb{E} \left[e^{-\beta \tilde{T}} G^0(X_{T_n^\Lambda}) I_{\{T_n^\Lambda < \infty\}} - \int_{T_{n-1}^\Lambda}^{T_n^\Lambda} e^{-\beta s} C(\Lambda_s) ds \middle| \mathcal{G}_{n-1} \right] - \int_0^{T_{n-1}^\Lambda} e^{-\beta s} C(\Lambda_s) ds \\ &\leq e^{-\beta T_{n-1}^\Lambda} G(X_{T_{n-1}^\Lambda}) - \int_0^{T_{n-1}^\Lambda} e^{-\beta s} C(\Lambda_s) ds \\ &\leq e^{-\beta T_{n-1}^\Lambda} G^0(X_{T_{n-1}^\Lambda}) - \int_0^{T_{n-1}^\Lambda} e^{-\beta s} C(\Lambda_s) ds = Y_{n-1}. \end{aligned}$$

□

Proposition 7. *Let G be an increasing, convex solution to (6.11) of at most linear growth. Then $H \leq G$.*

Proof. Let (τ, Λ) be any integrable strategy.

From Lemma 13 we have

$$\mathbb{E}[Y_1] = \mathbb{E} \left[e^{-\beta T_1^\Lambda} G^0(X_{T_1^\Lambda}) I_{\{T_1^\Lambda < \infty\}} - \int_0^{T_1^\Lambda} e^{-\beta s} C(\Lambda_s) ds \right] \leq G(x).$$

Moreover, since Y is a uniformly integrable supermartingale,

$$\begin{aligned} \mathbb{E}^x[Y_1] \geq \mathbb{E}^x[Y_\infty] &= \mathbb{E} \left[e^{-\beta \tau} G^0(X_\tau) I_{\{\tau < \infty\}} - \int_0^\tau e^{-\beta s} C(\Lambda_s) ds \right] \\ &\geq \mathbb{E} \left[e^{-\beta \tau} g(X_\tau) I_{\{\tau < \infty\}} - \int_0^\tau e^{-\beta s} C(\Lambda_s) ds \right]. \end{aligned}$$

Taking a supremum over stopping times and rate processes we conclude that $H(x) \leq G(x)$ (recall (6.2)) . \square

Our goal now is to show that $H = G$. We prove this result, first in the simplest case where the set of admissible rate processes is unrestricted (i.e. Λ_t takes values in $I = [0, \infty)$ and the cost function C is lower semi-continuous and convex, with $\lim_{\lambda \uparrow \infty} C(\lambda)/\lambda = \infty$). Then we argue that the same result holds true under weaker assumptions. Note that we allow for $\{\lambda \in I : C(\lambda) = \infty\}$ to be non-empty, but our assumption that C is lower semi-continuous means that if $\check{\lambda} = \inf\{\lambda : C(\lambda) = \infty\}$ then $C(\check{\lambda}) = \lim_{\lambda \uparrow \check{\lambda}} C(\lambda)$.

Theorem 10. *Suppose $I = [0, \infty)$ and $C : I \mapsto [0, \infty]$ is increasing, convex and lower semi-continuous with $\lim_{\lambda \uparrow \infty} C(\lambda)/\lambda = \infty$. Let G be an increasing, convex solution to (6.11) of at most linear growth. Then $H = G$.*

Proof. Let C' denote the right-derivative of C . Since $\lim_{\lambda \uparrow \infty} C(\lambda)/\lambda = \infty$, we set $C' = \infty$ on $\{\lambda : C(\lambda) = \infty\}$. Since C' is increasing it has a left-continuous inverse $D : \mathbb{R}_+ \mapsto \mathbb{R}_+$. In particular, $D(y) = \sup\{\lambda \in [0, \infty) : C'(\lambda) < y\}$ with the convention that $D(y) = 0$ if $C'(\lambda) \geq y$ on $(0, \infty)$. We note that our hypotheses mean that D is well defined and finite on $(0, \infty)$ and we set $D(0) = 0$.

Let $\hat{\Lambda} = (\hat{\Lambda}_s)_{s \geq 0}$ be given by $\hat{\Lambda}_s = D((g(X_s) - G(X_s))_+)$. We will show that $\hat{\Lambda}$ is the optimal rate process.

Note first that there is equality in (6.13), and therefore in (6.12), provided $F_s = F(g(X_s), G(X_s), \Lambda_s) = (g(X_s) - G(X_s))_+ \Lambda_s + \tilde{C}((g(X_s) - G(X_s))_+) = C(\Lambda_s)$. This is satisfied if $\Lambda_s = \hat{\Lambda}_s$.

Let $\mathcal{X}_> = \{x : g(x) > G(x)\}$ and let $\mathcal{X}_\leq = \{x : g(x) \leq G(x)\}$. Then, under the hypothesis of the theorem, whilst $X \in \mathcal{X}_\leq$ we have that $\hat{\Lambda} \equiv 0$. Hence (almost surely) $X_{T_1^{\hat{\Lambda}}} \in \mathcal{X}_>$ and $G^0(X_{T_1^{\hat{\Lambda}}}) = g(X_{T_1^{\hat{\Lambda}}})$. Then, taking $T = T_1^{\hat{\Lambda}}$ we have from

(6.12) that

$$\begin{aligned} G(x) &= \mathbb{E} \left[e^{-\beta T} G^0(X_T) I_{\{T < \infty\}} - \int_0^T e^{-\beta s} C(\Lambda_s) ds \right] \\ &= \mathbb{E} \left[e^{-\beta T} g(X_T) I_{\{T < \infty\}} - \int_0^T e^{-\beta s} C(\Lambda_s) ds \right] \leq H(x) \end{aligned}$$

and hence, combining with Proposition 7, $G = H$. \square

Corollary 5. $\Lambda^* = (\Lambda_s^*)_{s \geq 0}$ given by $\Lambda_s^* = D((g(X_s) - G(X_s))_+)$ is an optimal strategy where $D(y) = \sup\{\lambda \in [0, \infty) : C'(\lambda) < y\}$, and $\tau^* = T_1^{\Lambda^*}$ is an optimal stopping rule.

Our goal now is to extend Theorem 10 to allow for more general admissibility sets and cost functions.

Let c be a generic increasing, convex function $c : [0, \infty) \mapsto [0, \infty]$. If c takes the value $+\infty$ on $(\check{\lambda}, \infty)$ then we assume that $c(\check{\lambda}) = \lim_{\lambda \uparrow \infty} c(\lambda) = c(\infty)$, and set the right-derivative c' equal to infinity on $(\check{\lambda}, \infty)$ also. For such a c define $D_c : [0, \infty) \mapsto [0, \infty]$ by $D_c(y) = \sup\{\lambda \in (0, \infty) : c'(\lambda) < y\}$ again with the conventions that $D_c(y) = 0$ if $c'(\lambda) \geq y$ on $(0, \infty)$ and $D_c(0) = 0$. Note that $D_c(y) \leq \sup\{y : c(y) < \infty\}$.

Let I with endpoints $\{\underline{\lambda}, \bar{\lambda}\}$ be a subinterval of $[0, \infty)$ with the property that I is closed on the left and closed on the right if $\bar{\lambda} < \infty$.

Let $\gamma : I \mapsto \mathbb{R}_+$ be an increasing function. Let $\check{\gamma}$ be the largest convex minorant of γ on I . Then define γ^\dagger by $\gamma^\dagger(\lambda) = \gamma(\underline{\lambda})$ on $[0, \underline{\lambda})$ (if this interval is non-empty), $\gamma^\dagger(\lambda) = \check{\gamma}(\lambda)$ on $[\underline{\lambda}, \bar{\lambda}]$ and $\gamma^\dagger = \infty$ on $(\bar{\lambda}, \infty)$. By construction $\gamma^\dagger : [0, \infty) \mapsto [0, \infty]$ is convex and we can define D_{γ^\dagger} .

Suppose that $C : I \mapsto \mathbb{R}_+$ is our increasing, lower semi-continuous cost function. Introduce $C^\dagger : \mathbb{R}_+ \mapsto [0, \infty]$ and D_{C^\dagger} which we abbreviate to D^\dagger . Note that if $D^\dagger(z) < \underline{\lambda}$ then $z = 0$, $D^\dagger(z) = 0$ and $C^\dagger(0) = C^\dagger(\underline{\lambda}) = C(\underline{\lambda})$. Summarising the important results we have:

Lemma 15. $\tilde{C} = \widetilde{C^\dagger}$. Moreover, for $z \in [0, \infty)$, $C((D^\dagger(z) \vee \underline{\lambda}) \wedge \bar{\lambda}) = C^\dagger(D^\dagger(z))$.

Theorem 11. Suppose $I \subseteq [0, \infty)$ and let $C : I \mapsto \mathbb{R}$ be increasing, lower semi-continuous and such that $\lim_{\lambda \uparrow \infty} \frac{C(\lambda)}{\lambda} = \infty$. Let G be an increasing, convex solution of (6.11) and suppose G is of linear growth. Then $H = G$.

Proof. Introduce C^\dagger , defined from C as above, and let H^\dagger be the solution of the unrestricted problem (ie $I^\dagger = [0, \infty)$) with (convex) cost function C^\dagger . Note that

since $\tilde{C} = \tilde{C}^\dagger$ we have by Theorem 10 that $H^\dagger = G$. It remains to show that $H = H^\dagger$.

The inequality $H \leq H^\dagger$ is straight-forward: if (τ, Λ) is admissible for the interval I and integrable for cost function C , then it is admissible for the interval $[0, \infty)$ and integrable for cost function C^\dagger ; moreover $C \geq C^\dagger$, and so $H \leq H^\dagger$.

For the converse, let $\Lambda^\dagger = D^\dagger((g(X_s) - G(X_s))_+)$ and $\tau^\dagger = T_1^{\Lambda^\dagger}$ be optimal for the problem with cost function C^\dagger . Note that $\Lambda^\dagger \leq \bar{\lambda}$ and that

$$H^\dagger(x) = \mathbb{E}^x \left[e^{-\beta\tau^\dagger} g(X_{\tau^\dagger}) - \int_0^{\tau^\dagger} e^{-\beta s} C^\dagger(\Lambda_s^\dagger) ds \right]$$

Define $\Lambda^* = \underline{\lambda} \vee \Lambda^\dagger$ and $\tau^* = \tau^\dagger$. Then, by Lemma 15,

$$C(\Lambda_s^*) = C((D^\dagger((g(X_s) - G(X_s))_+) \vee \underline{\lambda}) \wedge \bar{\lambda}) = C^\dagger(D^\dagger((g(X_s) - G(X_s))_+)) = C^\dagger(\Lambda_s^\dagger).$$

Moreover, $\Lambda^* \in [\underline{\lambda}, \bar{\lambda}]$ and is admissible for the original problem with admissibility interval I . Then

$$H^\dagger(x) = \mathbb{E}^x \left[e^{-\beta\tau^*} g(X_{\tau^*}) - \int_0^{\tau^*} e^{-\beta s} C(\Lambda_s^*) ds \right] \leq H(x).$$

□

Remark 4. Note that $\Lambda^* \geq \Lambda^\dagger$ and we may have strict inequality if $\underline{\lambda} > 0$. In that case, when $g(X_s) \leq G(X_s)$ we have $\Lambda_s^\dagger = 0$, but $\Lambda_s^* = \underline{\lambda}$. In particular, we may have $\tau^* > T_1^{\Lambda^*}$, and the agent does not sell at the first opportunity. See Section 6.6.3.

6.5 Concave cost functions

In this section we provide a complementary result to Theorem 10 by considering a concave cost function C (defined on $\mathcal{I} = [0, \infty)$).

Suppose C is increasing and concave on $[0, \infty)$. Then the greatest convex minorant \check{C} of C is of the form

$$\check{C}(\lambda) = \delta + \epsilon\lambda$$

for some constants $\delta, \epsilon \in [0, \infty)$. Then, C and \check{C} have the same concave conjugates given by $\tilde{C}(z) := \inf_{\lambda > 0} \{C(\lambda) - \lambda z\}$ where $\tilde{C}(z) = \delta$ for $z \leq \epsilon$ and $\tilde{C}(z) = -\infty$ for $z > \epsilon$.

Recall that $g = (x - K)_+$. From the heuristics section we expect the value function

to solve (6.4). Then we might expect that on $g - H < \epsilon$ we have

$$\mathcal{L}^X H - \beta H - \delta = 0. \quad (6.14)$$

On the other hand some care is needed to interpret $\mathcal{L}^X H - \beta H = \tilde{C}((g - H)_+)$ on the set $g - H > \epsilon$. In fact, as we argue in the following theorem, $H \geq g - \epsilon$ and on the set $H = g - \epsilon$ (6.14) needs to be modified. We show that $H = w_{K,\epsilon,\delta}$ where (recall (6.3))

$$w_{K,\epsilon,\delta}(x) = \sup_{\tau \in \mathcal{T}([0,\infty))} \mathbb{E}^x \left[e^{-\beta\tau} \{ (X_\tau - K)_+ - \epsilon \} - \delta \int_0^\tau e^{-\beta s} ds \right]. \quad (6.15)$$

Recall from Section 6.2.1 that the continuation region and stopping region of $w_{K,\epsilon,\delta}$ are \mathcal{C} and \mathcal{S} respectively. The intuition is that when $H > g - \epsilon$ it is optimal to wait and to take $\Lambda = 0$ at cost δ per unit time on \mathcal{C} . However, on $H < g - \epsilon$ (and also when $H = g - \epsilon$) it is optimal to take Λ as large as possible on \mathcal{S} . Since there is no upper bound on Λ , this corresponds to taking Λ infinite — such a choice is inadmissible but can be approximated with ever larger finite values. Then, in the region where the agent wants to stop, if the stopping rate is large, say N , then the expected time to stop is N^{-1} , the cost incurred per unit time is $C(N) \approx \delta + \epsilon N$, and so the expected total cost of stopping is approximately $\frac{\delta + \epsilon N}{N} \approx \epsilon$. Effectively the agent can choose to sell (almost) instantaneously, for a fee or fixed transaction cost of ϵ . This explains why the problem value is the same as the problem value for (6.15).

Theorem 12. *Let $I = [0, \infty)$ and let $C : I \mapsto \mathbb{R}_+$ be non-negative, increasing and concave. Suppose the greatest convex minorant \check{C} of $C(\lambda)$ is of the form $\check{C}(\lambda) = \delta + \epsilon\lambda$ for non-negative constants δ and ϵ .*

Then $H(x) = w_{K,\epsilon,\delta}(x)$. The optimal strategy is to choose $\Lambda_s^ = 0$ when $X_s \in \mathcal{C}$; choose $\Lambda_s^* = \infty$ when $X_s \in \mathcal{S}$ and stop immediately.*

Proof. First we show that for any integrable τ and Λ

$$\mathbb{E}^x \left[e^{-\beta\tau} (X_\tau - K)_+ - \int_0^\tau e^{-\beta s} C(\Lambda_s) ds \right] \leq w_{K,\epsilon,\delta}(x).$$

Then we show that there is a sequence of admissible strategies for which the value function converges to this upper bound.

We prove the result in the case $\epsilon \geq \delta/\beta$ when the cost of taking $\Lambda = 0$ is small relative to the proportional cost $C(\lambda)/\lambda$ associated with taking Λ large. The proof in the case $\epsilon < \delta/\beta$ is similar, but slightly more complicated in certain verification

steps, because the explicit form of $w^{K,\epsilon,\delta}$ is not so tractable.

When $\epsilon \geq \delta/\beta$ we have that $w = w_{K,\epsilon,\delta}$ is given by

$$w(x) = \begin{cases} Ax^\theta - \frac{\delta}{\beta} & x \in \mathcal{C} \\ (x - K - \epsilon) & x \in \mathcal{S} \end{cases},$$

where $\mathcal{C} = (0, L)$ and $\mathcal{S} = [L, \infty)$, $L = \frac{\beta(K+\epsilon)-\delta}{\beta} \frac{\theta}{\theta-1}$ and $A = \frac{1}{\theta} L^{1-\theta}$. Let $w^0(x) = w(x) \vee (x - K)_+$. Note that since $\frac{\beta}{\mu} > \theta$ we have $\frac{\theta}{\theta-1} > \frac{\beta}{\beta-\mu}$ and $L > \frac{\beta(K+\epsilon)-\delta}{\beta-\mu}$.

For fixed Λ define $M^\Lambda = (M_t^\Lambda)_{t \geq 0}$ by

$$M_t^\Lambda = e^{-\int_0^t (\beta + \Lambda_s) ds} w(X_t) + \int_0^t e^{-\int_0^s (\beta + \Lambda_u) du} [\Lambda_s w^0(X_s) - C(\Lambda_s)] ds$$

and set $N_t = \int_0^t e^{-\int_0^s (\beta + \Lambda_u) du} \sigma X_s w'(X_s) dW_s$. Then $N = (N_t)_{t \geq 0}$ is a martingale and

$$dM_t^\Lambda = dN_t + e^{-\int_0^t (\beta + \Lambda_s) ds} [\mathcal{L}^X w - (\beta + \Lambda_t)w + \Lambda_t w^0 - C(\Lambda_t)] dt. \quad (6.16)$$

On $(0, L)$, $\mathcal{L}^X w - \beta w = \delta$, and (6.16) becomes

$$dM_t^\Lambda = dN_t + e^{-\int_0^t (\beta + \Lambda_s) ds} [\delta - \Lambda_t w + \Lambda_t w^0 - C(\Lambda_t)] dt \leq dN_t + e^{-\int_0^t (\beta + \Lambda_s) ds} [\Lambda_t (w^0 - w - \epsilon)] dt \leq dN_t.$$

since $w^0 \leq w + \epsilon$. Similarly, on (L, ∞) , $w(x) = (x - K) - \epsilon$ and since $L > K + \epsilon$, (6.16) yields

$$\begin{aligned} dM_t^\Lambda &\leq dN_t + e^{-\int_0^t (\beta + \Lambda_s) ds} [\mu X_t - (\beta + \Lambda)(X_t - K - \epsilon) + \Lambda_t(X_t - K) - (\delta + \epsilon \Lambda_t)] dt \\ &= dN_t + e^{-\int_0^t (\beta + \Lambda_s) ds} [(\mu - \beta)(X_t - L) + (\mu - \beta)L + \beta(K + \epsilon) - \delta] dt \leq dN_t. \end{aligned}$$

Putting the two cases together we see that M^Λ is a supermartingale for any strategy Λ .

The rest of the proof that $H \leq w$ follows exactly as in the the proofs of Lemma 13, Lemma 14 and Proposition 7, with w replacing G .

Now we show that there is a sequence of strategies for which the value function converges to $w = w_{K,\epsilon,\delta}$. Since $\delta + \epsilon\lambda$ is the largest convex minorant of C there exists $(\lambda_n)_{n \geq 1}$ with $\lambda_n \uparrow \infty$ such that $\frac{C(\lambda_n)}{\lambda_n} \rightarrow \epsilon$.

Consider first the strategy of a constant rate of search λ_n , with stopping at the first event time of the associated Poisson process. Let \tilde{H}_n denote the associated

value function. Then

$$\begin{aligned}
\tilde{H}_n(x) &= \mathbb{E}^x \left[\int_0^\infty \lambda_n e^{-\lambda_n t} dt \left\{ e^{-\beta t} (X_t - K)_+ - \int_0^t e^{-\beta s} C(\lambda_n) ds \right\} \right] \\
&\geq \int_0^\infty \lambda_n e^{-\lambda_n t} dt \left\{ e^{-\beta t} (x e^{\mu t} - K) - \int_0^t e^{-\beta s} C(\lambda_n) ds \right\} \\
&= \int_0^\infty \lambda_n e^{-(\lambda_n + \beta)t} (x e^{\mu t} - K) dt - \int_0^\infty e^{-\beta s} C(\lambda_n) ds \int_s^\infty \lambda_n e^{-\lambda_n t} dt \\
&= \frac{\lambda_n}{\lambda_n + \beta - \mu} x - \frac{\lambda_n}{\lambda_n + \beta} K - \frac{1}{\lambda_n + \beta} C(\lambda_n)
\end{aligned}$$

and $\tilde{H}_n(x) \rightarrow x - K - \epsilon$ as $n \uparrow \infty$. Suppose $\epsilon \geq \delta/\beta$. Let $L = \frac{\beta(K+\epsilon)-\delta}{\beta} \frac{\theta}{\theta-1}$ and let $\tau_L = \inf\{u : X_u \geq L\}$. Consider the strategy with rate $\hat{\Lambda}_n = \lambda_n I_{\{t \geq \tau_L\}}$, for which selling occurs at the first event time of the Poisson process with this rate, and let \hat{H}_n be the value function associated with this strategy.

For $x \geq L$ we have $\hat{H}_n(x) = \tilde{H}_n(x) \rightarrow x - K - \epsilon = w_{K,\epsilon,\delta}(x)$.

For $x < L$, we have $\mathbb{E}^x[e^{-\beta\tau_L}] = (\frac{x}{L})^\theta$ and

$$\begin{aligned}
\hat{H}_n(x) &= \mathbb{E}^x \left[e^{-\beta\tau_L} \tilde{H}_n(L) - \int_0^{\tau_L} e^{-\beta s} C(0) ds \right] \\
&= \mathbb{E}^x \left[e^{-\beta\tau_L} \left(\tilde{H}_n(L) + \frac{C(0)}{\beta} \right) - \frac{C(0)}{\beta} \right] \\
&= \left(\frac{x}{L} \right)^\theta \left[\tilde{H}_n(L) + \frac{\delta}{\beta} \right] - \frac{\delta}{\beta} \\
&\rightarrow w_{K,\epsilon,\beta}(x),
\end{aligned}$$

where the last line follows from the definition of L and some algebra. \square

6.5.1 An example

In this example we consider a cost function of the form $C(\lambda) = \sqrt{\lambda}$. Then a (plausibly) good strategy is to take $\Lambda_t = 0$ if $X_t < L^* = \frac{\theta}{\theta-1}$ and Λ_t very large otherwise. It is immediate that the value function H satisfies $H \leq w$; conversely, it is clear from Figure 6.11 that there exist strategies for which the value function is arbitrarily close to w .

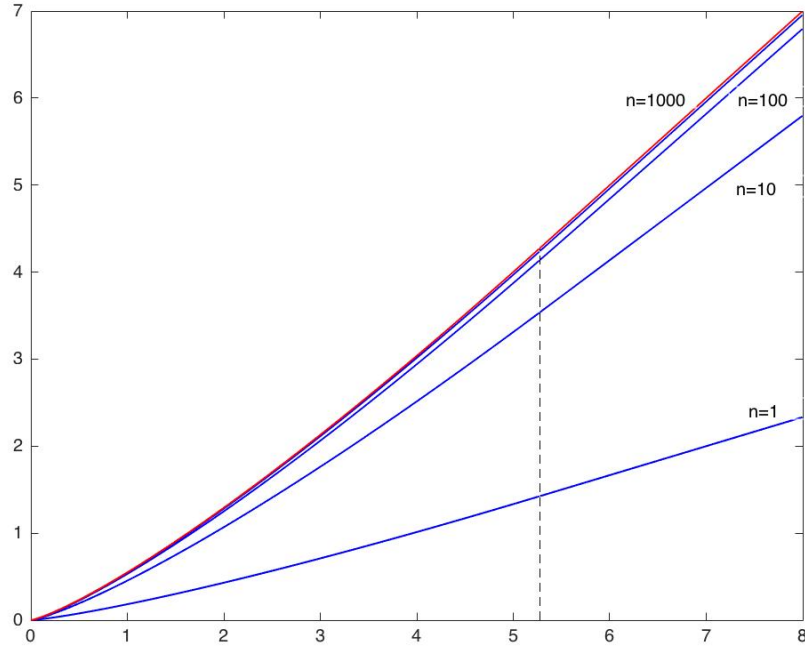


Figure 6.11: $(\beta, \mu, \sigma, K) = (5, 3, 3, 1)$; the highest line is $w = w_{K,0,0}$, and the other lines are the value functions under the rate function $\Lambda_n(x) = nI_{\{x \geq L^*\}}$.

6.6 Numerical analysis on further examples

6.6.1 Addition of a linear cost

Let C_0 be a convex, lower semi-continuous, increasing cost function, and consider the impact of adding a linear cost to C_0 ; in particular, let $C_b : \mathbb{R}_+ \mapsto \mathbb{R}_+$ be given by $C_b(\lambda) = C_0(\lambda) + \lambda b$ for $b > 0$.

Then the concave conjugates are such that $\tilde{C}_b(z) = \tilde{C}_0((z - b)_+)$.

Suppose further that G , the solution of (6.11) of linear growth, is such that $G \geq 0$ on \mathbb{R}_+ . The problem solution in the case of a purely quadratic cost function (recall Section 6.4.2) has this property. Then

$$(\{(x - K)_+ - G\}_+ - b)_+ = \{(x - (K + b))_+ - G\}_+.$$

It follows that

$$\tilde{C}_b(\{(x - K)_+ - G\}_+) = \tilde{C}_0((\{(x - K)_+ - G\}_+ - b)_+) = \tilde{C}_0(\{(x - (K + b))_+ - G\}_+)$$

and then that the value function for a payoff $(x - K)^+$ with cost function C_b is identical to the value function for a cost function $C_0(x)$ but with modified payoff $(x - (K + b))_+$.

Note that we see a similar result in the expansion (6.9) for G in the large x regime.

6.6.2 Quadratic costs with positive fixed cost

In this section we seek to generalise the results of Section 6.4.2 on purely quadratic cost functions to other quadratic cost functions. In view of the results in Section 6.6.1 the focus is on adding a positive intercept term, rather than a linear cost. Indeed the focus is on cost functions of the form $C(\lambda) = a + \frac{c}{2}\lambda^2$ for $a > 0$.

In this section we will take a and c fixed and compare the cost functions $C_0(\lambda) = \frac{c}{2}\lambda^2$, $C_1(\lambda) = a + \frac{c}{2}\lambda^2$ and $C_>(\lambda) = aI_{\{\lambda > 0\}} + \frac{c}{2}\lambda^2$. The difference between the last two cases is that in the final case, not searching at all incurs zero cost, whereas in the middle case, there is a fixed cost which applies irrespective of whether there is a positive rate of searching for offers or not.

In Section 6.4.2 we saw that H_0 , the value function for the cost $C_0(\lambda) = \frac{c}{2}\lambda^2$, solves

$$\mathcal{L}^X H_0 - \beta H_0 = \frac{[(g - H_0)_+]^2}{2c}.$$

There is a threshold L with $L > K$, such that $H_0 > g$ on $(0, L)$ and $H_0 < g$ on (L, ∞) . On $(0, L)$ we have that $H_0(x) = (L - K)\frac{x^\theta}{L^\theta}$; on $[L, \infty)$, H_0 solves $\frac{1}{2}\sigma^2 x^2 h'' + \mu x h' - \beta h = \frac{1}{2c}(x - K - h)^2$ subject to initial conditions $H_0(L) = (L - K)$ and $H'_0(L) = \theta \frac{L - K}{L}$. We adjust L until we find a solution for which H_0 is of linear growth at infinity.

Now consider C_1 with associated value function H_1 . When X is very small, there is little prospect of X ever rising above K . Nonetheless the agent faces a fixed cost, even if she does not search for offers. It will be cheaper to search for offers, because although the payoff is zero when a candidate purchaser is found, it is then possible in our model to stop paying the fixed cost.

Suppose $X = 0$. If the agent chooses to search for buyers at rate λ then the expected time until a buyer is found is λ^{-1} . The expected discounted cost until a buyer is found is

$$\int_0^\infty \lambda e^{-\lambda s} \int_0^s e^{-\beta u} \left(a + \frac{c}{2}\lambda^2 \right) du = \frac{a + \frac{c}{2}\lambda^2}{\beta + \lambda}.$$

This is minimised by the choice $\lambda = \lambda_*$ where $\lambda_* = \sqrt{\beta^2 + \frac{2a}{c}} - \beta$ and the minimal

cost is h_*^- where

$$h_*^- = \frac{a + \frac{c}{2}\lambda_*^2}{\beta + \lambda_*} = c\lambda_* = c \left[\sqrt{\beta^2 + \frac{2a}{c}} - \beta \right].$$

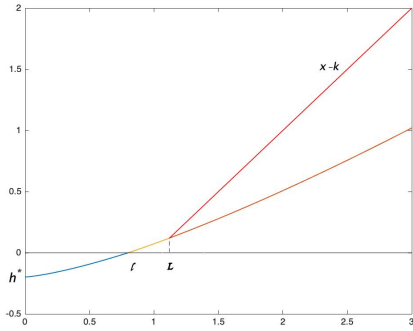
Then $H_1(0) = -h_*^-$. (Another way to see this is to note that at 0 we expect $\mathcal{L}^X H_1 = 0$ and therefore $H_1(0)$ to solve $-\beta h = \tilde{C}(-h) = a - \frac{h^2}{2c}$.) Then, the value function H_1 is such that there exists ℓ and L with $0 < \ell < K < L < \infty$ such that H_1 is C^1 with $H_1 < 0$ on $(0, \ell)$, $H_1(x) > (x - K)_+$ on (ℓ, L) and $H_1(x) < (x - K)_+$ on (L, ∞) and such that H_1 satisfies

$$\mathcal{L}^X h - \beta h = \begin{cases} a - \frac{1}{2c} h^2 & x < \ell; \\ a & \ell < x < L; \\ a - \frac{1}{2c} (g - h)^2 & L < x. \end{cases}$$

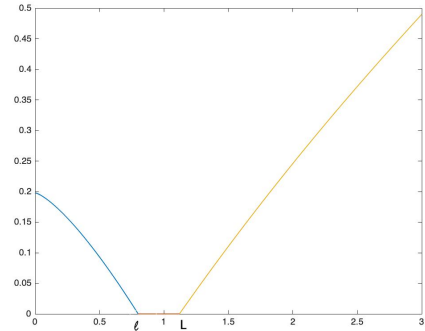
See Figure 6.12. Considering H_1 on (ℓ, L) we have $H_1(x) = Ax^\theta + Bx^\phi - \frac{a}{\beta}$ for some constants A and B chosen so that $H_1(\ell) = 0$ and $H_1(L) = (L - K)$:

$$A = \frac{L^{-\phi}(L - K + \frac{a}{\beta}) - \ell^{-\phi}\frac{a}{\beta}}{L^{\theta-\phi} - \ell^{\theta-\phi}}, \quad B = \frac{\ell^{-\phi}L^{\theta-\phi}\frac{a}{\beta} - \ell^{\theta-\phi}L^{-\phi}(L - K + \frac{a}{\beta})}{L^{\theta-\phi} - \ell^{\theta-\phi}}.$$

Then for general ℓ and L we can use value matching and smooth fit at ℓ and L to construct a solution on $(0, \infty)$. Finally, we adjust ℓ and L until $H_1(0) = -h_*^-$ and H_1 has linear growth.



(a) The value function $H_1(x)$.



(b) The optimal rate $\Lambda_1^*(x)$.

Figure 6.12: $(\beta, \mu, \sigma, K) = (5, 3, 2, 1)$. The cost function is $C_1(\lambda) = 1 + \lambda^2$. The left figure shows the value function, and the right figure the optimal stopping rate. There are two critical thresholds $\ell = \ell^*$ and $L = L^*$.

In Figure 6.12 we plot the value function and optimal rate for the Poisson process

for $C_1(\lambda) = 1 + \lambda^2$. There are two critical thresholds ℓ^* and L^* with $0 < \ell^* < K < L^*$. Above L^* the agent would like to stop in order to receive the payoff $(x - K)$, and is willing to expend effort to try to generate selling opportunities in order to receive the payoff before discounting reduces the worth. Below ℓ^* the agent would like to stop, even though the payoff is zero, and is willing to expend effort to generate stopping opportunities in order to limit the costs they incur prior to stopping. Between ℓ^* and L^* the agent does not expend any effort searching for offers and would not accept any offers which were received.

Now consider the cost function $C_{>}(\lambda) = aI_{\{\lambda>0\}} + \frac{\epsilon}{2}\lambda^2$ with associated value function $H_{>}$. We have $\tilde{C}_{>}(z) = 0$ for $z \leq \sqrt{2ac}$ and $\tilde{C}_{>} = a - \frac{z^2}{2c}$ for $z \geq \sqrt{2ac}$. As in the pure quadratic case, there is always the option of taking $\Lambda \equiv 0$ at zero cost, so that the value function is non-negative. It follows that $H_{>}(0) = 0$. There is a threshold L below which the agent does not search for offers. But, this threshold is not the boundary between the sets $\{x : H_{>}(x) > g(x)\}$ and $\{x : H_{>}(x) < g(x)\}$, since when $g(x) - H_{>}(x)$ is small, it is still preferable to take $\Lambda = 0$, rather than to incur the cost of strictly positive λ . Instead L separates the sets $\{x : H_{>}(x) > g(x) - \sqrt{2ac}\}$ and $\{x : H_{>}(x) < g(x) - \sqrt{2ac}\}$. We find that there is a threshold L with $L > K$ such that on $(0, L)$, $H_{>}$ solves $\mathcal{L}^X h - \beta h = a$. At L we have $H_{>}(L) = (L - K - \sqrt{2ac})$ and it follows that on $(0, L)$ we have $H_{>}(x) = \frac{L - K - \sqrt{2ac}}{L^\theta} x^\theta$. Then, on (L, ∞) , $H_{>}$ solves $\mathcal{L}^X h - \beta h = a - \frac{(x - K - h)^2}{2c}$, subject to value matching and smooth fit conditions at $x = L$. Finally, we adjust the value of the threshold L until H is of linear growth for large x .

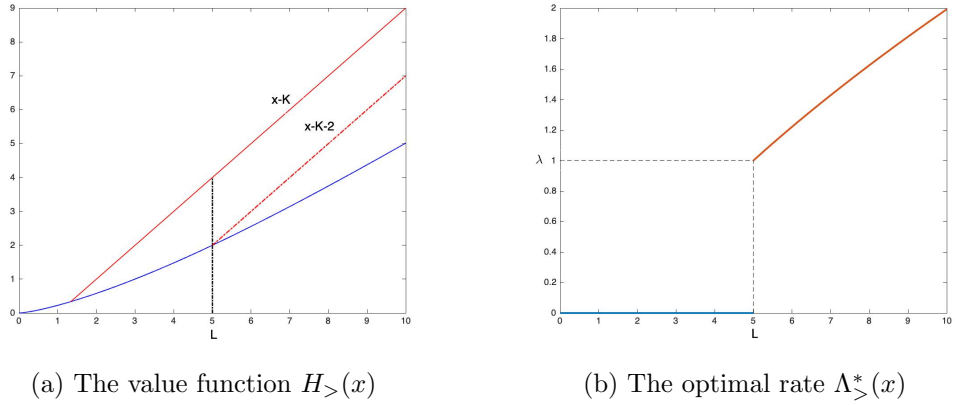
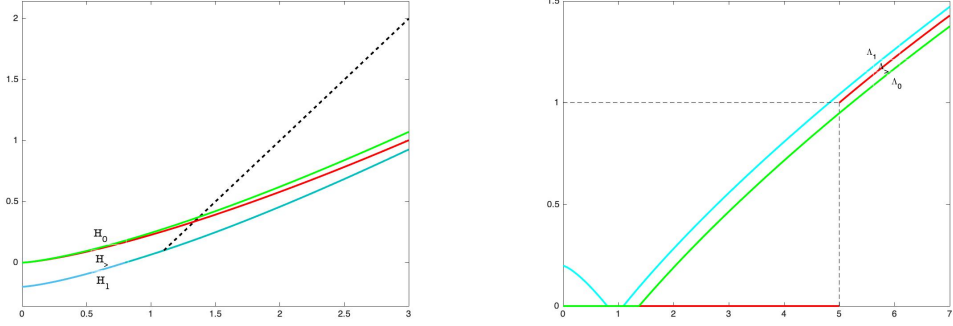


Figure 6.13: $(\beta, \mu, \sigma, K) = (5, 3, 2, 1)$. The cost function is $C_{>}(\lambda) = I_{\{\lambda>0\}} + \lambda^2$. The highest convex minorant is $\tilde{C}_{>}(\lambda) = \lambda + [(\lambda - 1)_+]^2$. (Here we use the fact that $\sqrt{2ac} = 2$.)

In Figure 6.13 we plot the value function $H_{>}$ and optimal rate $\Lambda_{>}^*$. We see that

$\Lambda_{>}^*$ never takes values in $(0, 1)$ where $C_{>} > \check{C}_{>}$. Either it is optimal to spend a non-negligible amount of effort on searching for candidate buyers, or it is optimal to spend no effort.



(a) Comparison of the value functions (b) Comparison of the optimal stopping rates

Figure 6.14: $(\beta, \mu, \sigma) = (5, 3, 2, 1)$. The cost functions we consider are $C_0(\lambda) = \lambda^2$, $C_{>}(\lambda) = I_{\{\lambda > 0\}} + \lambda^2$ and $C_1(\lambda) = 1 + \lambda^2$. The left figure plots the value functions under optimal behaviour, and the right figure plots the optimal rates for the Poisson process. For $x > 5$ we have $\Lambda_1^* > \Lambda_{>}^* > \Lambda_0^*$. For small x , $\Lambda_1^* > 0 = \Lambda_{>}^* = \Lambda_0^*$.

Figure 6.14 compares the value functions and optimal rates for the Poisson process for the three cost functions $C_0(\lambda) = \lambda^2$, $C_{>}(\lambda) = I_{\{\lambda > 0\}} + \lambda^2$ and $C_1(\lambda) = 1 + \lambda^2$. Since $C_0 \leq C_{>} \leq C_1$ we must have that $H_0 \geq H_{>} \geq H_1$ and we see that away from $x = 0$ this inequality is strict. Indeed, especially for small x , H_0 and $H_{>}$ are close in value. The differences in optimal strategies are more marked. For large x the fact that $H_0 > H_{>} > H_1$ means that $\Lambda_0^* < \Lambda_{>}^* < \Lambda_1^*$, and thus that even though $C_1 > C_0$, the agent searches at a higher rate under C_1 than under C_0 . Note that, we only have $\Lambda_{>}^* > 0$ for x above a critical value (in our case, approximately 5). Conversely, for C_1 there is a second region where $\Lambda_1 > 0$, namely where x is small.

6.6.3 Cost functions defined on a subset of \mathbb{R}_+

In this section we consider the case where there is a strictly positive lower bound on the rate at which offers are received. In fact, in our example the optimal rate of offers takes values in a two-point set. Nonetheless, we see a rich range of behaviours.

Suppose Λ takes values in $[\underline{\lambda}, \bar{\lambda}]$ where $0 < \underline{\lambda} < \bar{\lambda} < \infty$ and suppose $C : [\underline{\lambda}, \bar{\lambda}] \mapsto \mathbb{R}_+$ is increasing and concave. Introduce $\check{C} : [\underline{\lambda}, \bar{\lambda}] \mapsto [0, \infty)$ defined by $\check{C}(\lambda) =$

$C(\underline{\lambda}) + \frac{\lambda - \underline{\lambda}}{\bar{\lambda} - \underline{\lambda}}(C(\bar{\lambda}) - C(\underline{\lambda}))$. Finally introduce $C^\dagger : [0, \infty) \mapsto [0, \infty]$ by

$$C^\dagger(\lambda) = \begin{cases} C(\underline{\lambda}) & \lambda < \underline{\lambda}, \\ \check{C}(\lambda) & \underline{\lambda} \leq \lambda \leq \bar{\lambda}, \\ \infty & \bar{\lambda} < \lambda. \end{cases}$$

Write $a = C(\underline{\lambda})$ and $b = \frac{C(\bar{\lambda}) - C(\underline{\lambda})}{\bar{\lambda} - \underline{\lambda}}$. Then C^\dagger has concave conjugate $\tilde{C}^\dagger(z) = a - \underline{\lambda}z$ for $z \leq b$ and $\tilde{C}^\dagger(z) = a - b\underline{\lambda} - (z - b)\bar{\lambda}$ for $z > b$.

Suppose first that $C(\underline{\lambda}) = a = 0$. Then the value function H is positive, increasing and C^1 and satisfies

$$\mathcal{L}^X h - \beta h = \begin{cases} 0 & x < L, \\ -\underline{\lambda}(g - h) & L \leq x \leq M, \\ -b\underline{\lambda} - \bar{\lambda}(g - h - b) & M < x, \end{cases}$$

where L and M are constants satisfying $0 < K < L < M$ which must be found as part of the solution, and are such that $h(x) > (x - K)$ on $(0, L)$, $(x - K) > h(x) > x - K - b$ on (L, M) and $(x - K - b) > h(x)$ on (M, ∞) . See Figure 6.15.

Fix L and consider constructing a solution to the above problem with $H(0)$ bounded. On $(0, L)$ we have that $H(x) = Ax^\theta + Bx^\phi$ and the requirement that H is bounded means that $B = 0$ and then $A = (L - K)L^{-\theta}$. We then use the C^1 continuity of H at L to find the constants C and D in the expression for H over (L, M) :

$$H(x) = Cx^\theta + Dx^\phi + \frac{\lambda}{\underline{\lambda} + \beta - \mu}x - \frac{K\underline{\lambda}}{\underline{\lambda} + \beta}. \quad (6.17)$$

Here ϕ, θ with $\phi < 0 < 1 < \theta$ are solutions to $Q_\lambda(\cdot) = 0$ where $Q_\lambda(\psi) = \frac{1}{2}\sigma^2\psi(\psi - 1) + \mu\psi - (\beta + \lambda)$. In turn, we can find the value of $M = M(L)$ where H given by (6.17) crosses the line $y(x) = x - K - b$, and then value matching at M gives us the value of E in the expression for H over $[M, \infty)$:

$$H(x) = Ex^{\bar{\phi}} + \frac{\bar{\lambda}}{\bar{\lambda} + \beta - \mu}x - \frac{(K + b)\bar{\lambda} - b\underline{\lambda}}{\bar{\lambda} + \beta}$$

where $\bar{\phi}$ is the negative root of $Q_{\bar{\lambda}}(\cdot) = 0$. (There is no term of the form $x^{\bar{\theta}}$ since H must be of linear growth at infinity.) Finally, we can solve for L by matching derivatives of H at M .

Figure 6.15 plots the value function and the optimal rate function. The state space splits into three regions. On $x > M$ the asset is considerably in-the-money and the agent is prepared to pay the cost to generate a higher rate of selling opportunities.

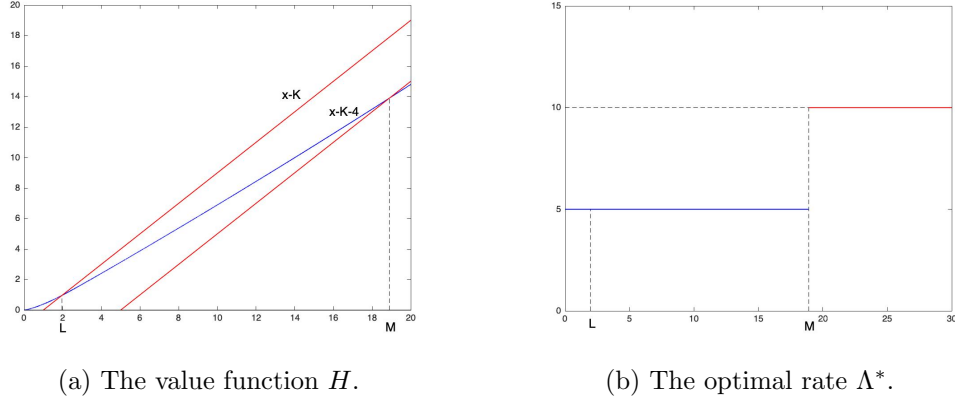


Figure 6.15: $(\beta, \mu, \sigma, K, \underline{\lambda}, \bar{\lambda}, C(\underline{\lambda}), C(\bar{\lambda})) = (5, 3, 2, 1, 5, 10, 0, 20)$. Note that $b = \frac{C(\bar{\lambda}) - C(\underline{\lambda})}{\bar{\lambda} - \underline{\lambda}} = 4$. The left figure plots the value function and the right figure plots the optimal rate function. Λ is constrained to lie in $[5, 10]$, and the cost function is $20I_{\{\lambda > 5\}}$. We see that Λ^* takes values in $\{5, 10\}$.

When x is not quite so large, and $L < x < M$, the agent is not prepared to pay this extra cost, but will sell if opportunities arise. However, if $x < L$ then selling opportunities will arise (we must have $\Lambda \geq \underline{\lambda}$) but the agent will forgo them. Ideally the agent would choose $\Lambda = 0$, but this is not possible. Instead the agent takes $\Lambda = \underline{\lambda}$, but synthesises a rate of zero, by rejecting all offers.

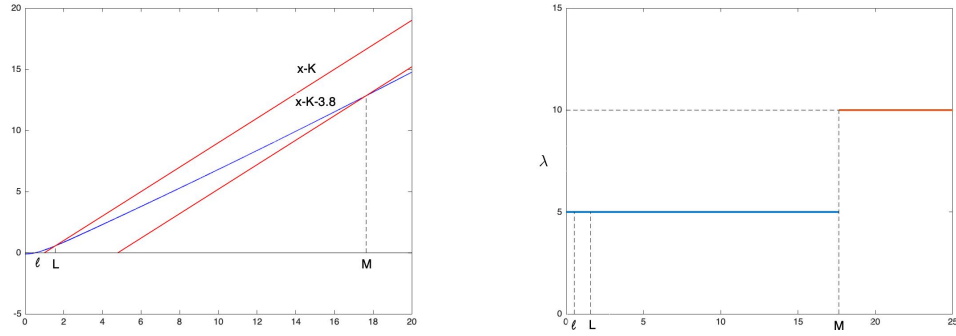
When $C(\underline{\lambda}) > 0$, the agent will not pay the fixed cost indefinitely when X is small. The behaviour for large X is unchanged, but the agent will now stop if offers arrive when the value of continuing is negative, including when X is near zero. There are two cases depending on whether $\frac{C(\underline{\lambda})}{\underline{\lambda} + \beta} \leq \frac{C(\bar{\lambda})}{\bar{\lambda} + \beta}$ or otherwise. In the former case, when X is small it is cheaper to pay the lower cost and to stop if opportunities arise, than to pay the higher cost with the hope of stopping sooner. In the latter case, the comparison is reversed. We find that H solves

$$\mathcal{L}^X h - \beta h = \tilde{C}((g - h)_+)$$

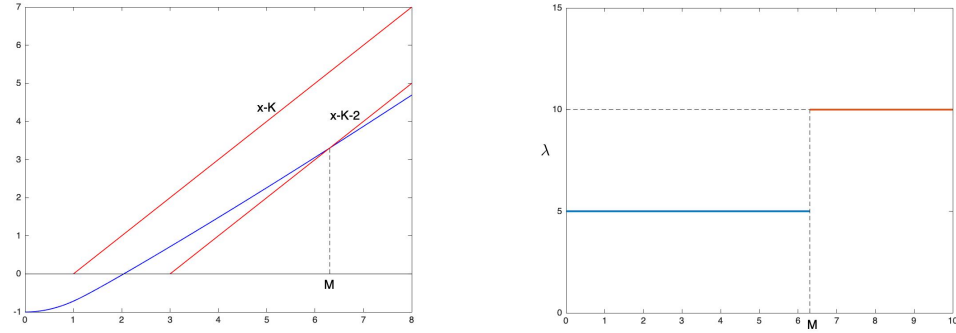
subject to $h(0) = -\min_{\lambda \in \{\underline{\lambda}, \bar{\lambda}\}} \{\frac{C(\lambda)}{\lambda + \beta}\}$ and the fact that h is of linear growth at infinity. The solution is smooth, except at points where $\tilde{C}((g - h)_+)$ is not differentiable. This may be at K where g is not differentiable, or when $g = h$, or, since \tilde{C} is non-differentiable at b , when $g - h = b$.

Figure 6.16 shows the value function and the optimal search rate in the case where $\frac{C(\underline{\lambda})}{\underline{\lambda} + \beta} \leq \frac{C(\bar{\lambda})}{\bar{\lambda} + \beta}$. This means that when x is small the agent expends as little effort as possible searching for offers, although they do accept any offers which arrive. There

is also a critical threshold M , beyond which it is optimal to put maximum effort into searching for offers. There are then two sub-cases depending on whether costs are small or large. If costs are large then the agent will always accept any offer which comes along (Figure 6.16(c) and (d)). However, when costs are small (Figure 6.16(a) and (b)), there is a region (ℓ, L) over which $h(x) > g(x) = (x - K)_+$. Then, as in the region $(0, L)$ when $C(\underline{\lambda}) = 0$, even when there is an offer the agent chooses to reject it. Effectively, the agent creates a zero rate of offers by thinning out all the events of the Poisson process.



(a) The value function H in the case $(C(\underline{\lambda}) = 1, C(\bar{\lambda}) = 20)$. (b) The optimal rate Λ^* in the case $(C(\underline{\lambda}) = 1, C(\bar{\lambda}) = 20)$.



(c) The value function H in the case $(C(\underline{\lambda}) = 10, C(\bar{\lambda}) = 20)$. (d) The optimal rate Λ^* in the case $(C(\underline{\lambda}) = 10, C(\bar{\lambda}) = 20)$.

Figure 6.16: $(\beta, \mu, \sigma, K, \underline{\lambda}, \bar{\lambda}) = (5, 3, 2, 1, 5, 10)$. The left panels plot the value function and the right panels plot the optimal rate function. In each row $\frac{C(\underline{\lambda})}{\underline{\lambda} + \beta} < \frac{C(\bar{\lambda})}{\bar{\lambda} + \beta}$. In the case of lower costs ($C(\underline{\lambda}) = 1$) there is a region (ℓ, L) where $H(x) > g(x)$ and the agent chooses to continue rather than to stop.

Figure 6.17 shows the value function and the optimal search rate in the case where $\frac{C(\underline{\lambda})}{\underline{\lambda} + \beta} > \frac{C(\bar{\lambda})}{\bar{\lambda} + \beta}$. Then, necessarily, $b = \frac{C(\bar{\lambda}) - C(\underline{\lambda})}{\bar{\lambda} - \underline{\lambda}} < \frac{C(\bar{\lambda})}{\bar{\lambda} + \beta}$. When x is small the agent searches at the maximum rate to generate an offer as quickly as possible.

Necessarily $H(0) < -b$. If costs are large enough, then $H(x) < (x - k)_+ - b$ for all x , see Figure 6.17(a) and (b). Then the agent wants to stop as soon as possible, and is prepared to pay the higher cost rate in order to facilitate this. As costs decrease, we may have $(x - k)_+ - b \leq H(x)$ for some x , whilst the inequality $H(x) < (x - k)_+$ remains true, see Figure 6.17(c) and (d). Then there is a region (m, M) over which the optimal strategy is $\Lambda^*(x) = \underline{\lambda}$. The agent still accepts any offer which is made. Finally, if costs are small enough we find that there is a neighbourhood (ℓ, L) of K for which $H(x) > (x - K)_+$. Then, on (ℓ, L) the agent takes $\Lambda^*(x) = \underline{\lambda}$, but chooses to continue rather than stop if any offers are made.

As a limiting special case suppose $\bar{\lambda} = \underline{\lambda} = \hat{\lambda}$ and that $C(\hat{\lambda}) = c \in [0, \infty)$. Then there is a single threshold L to be determined and H is of the form

$$H(x) = \begin{cases} Ax^\theta - \frac{c}{\beta} & x \leq L \\ Bx^{\hat{\phi}} + \frac{\hat{\lambda}}{\hat{\lambda} + \beta - \mu}x - \frac{(c + \hat{\lambda}K)}{\beta + \hat{\lambda}} & x > L \end{cases}$$

where $\hat{\phi}$ is the negative root of $Q_{\hat{\lambda}}(\cdot) = 0$. The value matching condition $H(L) = (L - K)_+$ gives that $A = L^{-\theta}(L - K + \frac{c}{\beta})$ and

$$B = L^{-\hat{\phi}} \left\{ \left(\frac{\beta - \mu}{\hat{\lambda} + \beta - \mu} \right) L + \frac{c - \beta K}{\beta + \hat{\lambda}} \right\}.$$

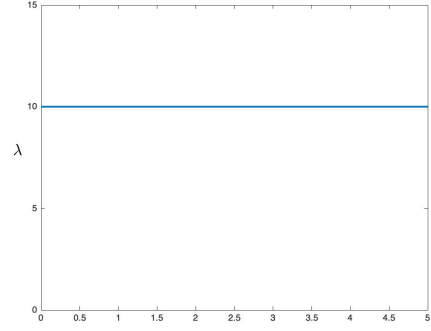
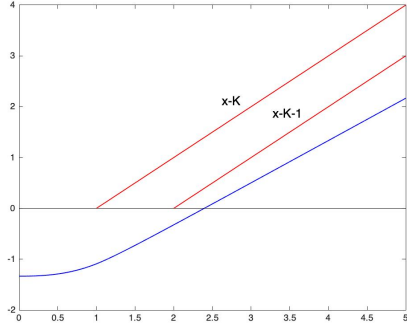
Then first order smooth fit at L implies that

$$L = (\beta K - c) \left[\frac{\theta}{\beta} - \frac{\hat{\phi}}{\beta + \hat{\lambda}} \right] \left\{ \theta - \frac{\hat{\phi}(\beta - \mu)}{\hat{\lambda} + \beta - \mu} - \frac{\hat{\lambda}}{\hat{\lambda} + \beta - \mu} \right\}^{-1}.$$

Note that if we take $c = 0$ we recover exactly the expressions in Dupuis and Wang (2005).

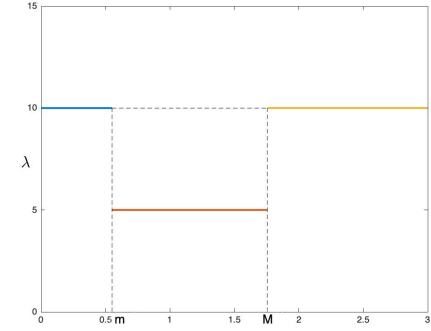
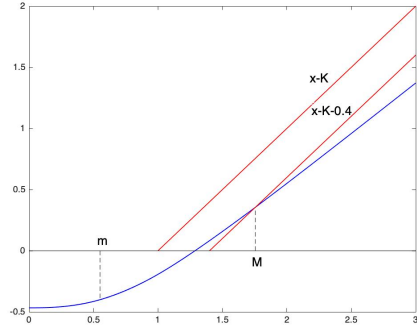
6.7 Conclusion

Our goal in this article is to extend the analysis of Dupuis and Wang (2005) who considered optimal stopping problems where the stopping time was constrained to lie in the event times of a Poisson process, to allow the agent to affect the frequency of those event times. The motivation was to model a form of illiquidity in trading and to consider problems in which the agent can exert effort in order to increase the opportunity set of candidate moments when the problem can terminate. This notion of effort is different to the idea in the financial economics literature of managers expending effort in order to change the dynamics of the underlying process, as



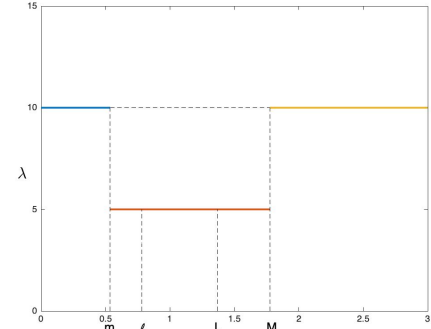
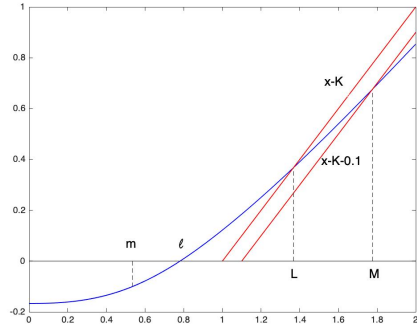
(a) The value function H in the case $(C(\underline{\lambda}) = 15, C(\bar{\lambda}) = 20)$. $b = 1$.

(b) The optimal rate Λ^* in the case $(C(\underline{\lambda}) = 15, C(\bar{\lambda}) = 20)$.



(c) The value function H in the case $(C(\underline{\lambda}) = 5, C(\bar{\lambda}) = 7)$. $b = 0.4$.

(d) The optimal rate Λ^* in the case $(C(\underline{\lambda}) = 5, C(\bar{\lambda}) = 7)$.



(e) The value function H in the case $(C(\underline{\lambda}) = 2, C(\bar{\lambda}) = 2.5)$. $b = 0.1$.

(f) The optimal rate Λ^* in the case $(C(\underline{\lambda}) = 2, C(\bar{\lambda}) = 2.5)$.

Figure 6.17: $(\beta, \mu, \sigma, K, \underline{\lambda}, \bar{\lambda}) = (5, 3, 2, 1, 5, 10)$. The left column plots the value function and the right column plots the optimal rate function. In each row $\frac{C(\underline{\lambda})}{\underline{\lambda} + \beta} > \frac{C(\bar{\lambda})}{\bar{\lambda} + \beta}$. Near $x = 0$ it is always preferable to choose the maximum possible rate process. Costs decrease as we move down the rows.

exemplified by Sannikov (2008) but seems appropriate for the context.

Our work focuses on optimal stopping of an exponential Brownian motion under a perpetual call-style payoff, although it is clear given the work of Lempa (2007) how the analysis could be extended to other diffusion processes and other payoff functions. Nonetheless, even in this specific case we show how it is possible to generate a rich range of possible behaviours, depending on the choice of cost function. In our time-homogeneous, Markovian set-up, the rate of the Poisson process can be considered as a proxy for effort, and the problem can be cast in terms of this control variable. Then, the form of the solution depends crucially on the shape of the cost function, as a function of the rate of the inhomogeneous Poisson process.

One important quantity is the limiting value for large λ of the average cost $\frac{C(\lambda)}{\lambda}$. If this limit is infinite, then the agent does not want to select very large rates for the Poisson process as they are too expensive. In this case we can replace C with its convex minorant and solve the problem for that cost function. However, if C is concave and the set of possible values for the rate process is unbounded then when the asset is sufficiently in the money, the agent wants to choose an infinite rate function, and thus to generate a stopping opportunity immediately. Choosing a very large rate function, albeit for a short time, incurs a cost equivalent to a fixed fee for stopping, and this is reflected in the form of the value function.

Another important quantity is the value of C at zero. If a choice of zero stopping rate is feasible and incurs zero cost per unit time, then the agent always has a feasible, costless choice for the rate function, and the value function is non-negative. Then, when the asset price is close to zero we expect the agent to put no effort into searching for buyers, and to wait. However, if the cost of choosing a zero rate for the Poisson process is strictly positive, then the agent has an incentive to search for offers even when the asset price is small and the payoff is zero. When the agent receives an offer they accept, because this ends their obligation to pay costs. In this way we can have a range of optimal behaviours when the asset price is small.

When the range of possible rate processes includes zero and C is strictly increasing, then the agent only exerts effort to generate selling opportunities in circumstances where they would accept those opportunities. The result is that the agent stops at the first event of the Poisson process, and the optimal stopping element of the problem is trivial. However, an interesting feature arises when there is a lower bound on the admissible rate process. Then, the agent may receive unwanted offers, which they choose to decline. In this case the agent chooses whether to accept the first offer or to continue.

We model the cost function C as increasing, which seems a natural requirement of

the problem. (However, if C is not increasing, we can introduce a largest increasing cost function which lies below C , and the value function for that problem will match the solution of the original problem.) We also assume that the interval of possible values for the rate process is closed (at any finite endpoints) and that C is lower semi-continuous. Neither of these assumptions is essential although they do simplify the analysis. In particular, these assumptions ensure that the minimal cost is attained, and that we do not need to consider a sequence of approximating strategies and problems.

Bibliography

- [1] Agranov, M. and P. Ortoleva, (2017), Stochastic choice and preferences for randomisation, *Journal of Political Economy*, 125(1), 40-68.
- [2] Ang, A. D. Papanikolaou and M.M. Westerfield, (2014), Portfolio choice with illiquid assets, *Management Science*, 60(11), p2381–2617.
- [3] Azéma J. and M. Yor, (1979), Une solution simple au problème de Skorokhod, *Sem. de Prob. XIII*, 90-115.
- [4] Barberis, N., (2012), A Model of Casino Gambling, *Management Science*, 58, 35-51.
- [5] Camerer, C. F., and T. Ho, (1994), Violations of the betweenness axiom and nonlinearity in probabilities, *Journal of Risk and Uncertainty*, 8, 167-196.
- [6] Cerreia-Vioglio, S., D.Dillenberger, and P. Ortoleva, (2015), Cautious Expected Utility and the Certainty Effect, *Econometrica*, 83, 693-728
- [7] Cerreia-Vioglio, S., D.Dillenberger, P. Ortoleva, and G. Riella, (2017), Deliberately Stochastic, Working paper, Columbia University.
- [8] Cox, A.M.G. and D.G. Hobson, (2004), An optimal Skorokhod embedding for diffusions, *Stochastic Processes and Applications*, 111, 17-39.
- [9] Dayanik, S. and I. Karatzas, (2003), On the optimal stopping problem for one-dimensional diffusions, *Stoc. Proc. & Appl*, 107, 2, 173-212.
- [10] Dillenberger, D., (2010), Preference of one-shot resolution of uncertainty and Allais -type behaviour, *Deliberately Stochastic*. Manuscript, Columbia University.
- [11] Dixit, A.K. and R.S. Pindyck, (1994), *Investment under uncertainty*, Princeton, NJ.

- [12] Dupis, P. and H. Wang, (2005), Optimal Stopping with Random Intervention Times, *Advances in Applied Probabilities*, 34(1), 141-157
- [13] Dwenger N., Kübler D., and G. Weizsäcker, (2013), Flipping a Coin: Evidence from Laboratory and Field, *SSRN Working paper*, id2353282.
- [14] Durrett, R., (1991), *Probability: Theory and Examples*, Wadsworth, Pacific Grove, California.
- [15] Ebert, S. and P. Strack, (2015), Until the Bitter End: On Prospect Theory in a Dynamic Context, *American Economic Review*, 105(4), 1618-1633.
- [16] Fudenberg, D., R. Iijima and T. Strzalecki, (2015), Stochastic choice and revealed perturbed utility, *Econometrica*. 83 (6): 2371-2409.
- [17] Fudenberg, D., P. Strack and T. Strzalecki, (2017), Speed, Accuracy and the Optimal Timing of Choices, Working paper, Harvard University.
- [18] Gul, F. ,and W. Pesendorfer, (2006), Random Expected Utility, *Econometrica*, 74,121-146.
- [19] Guo, X and J. Liu, (2005), Stopping at the maximum of geometric Brownian motion when signals are received, *J. Appl. Probab.*, 43(3) p826-838.
- [20] Hall, W.J., (1998), On the Skorokhod embedding theorem, *Technical Report 33*, Stanford University, Department of Statistics.
- [21] Henderson, V., D. Hobson, and A.S.L.Tse, (2017), Randomised Strategies and Prospect Theory in a Dynamic Context, *Journal of Economic Theory*, 168, 287-300.
- [22] Henderson, V., D. Hobson, and M. Zeng, (2017), Optimal Stopping and the Sufficiency of randomised Thresholds, Working paper.
- [23] Henderson, V., D. Hobson, and M. Zeng, (2017), Cautious Stochastic Choice, Optimal Stopping and Deliberate randomisation, Working paper, University of Warwick.
- [24] Henderson, V., (2012), Prospect Theory, Liquidation and the Disposition Effect, *Management Science*, 58, 2, February 2012, 445-460.
- [25] He, X., S. Hu, J. Obloj and X.Y. Zhou, (2017), Path dependent and randomised strategies in Barberis' Casino Gambling model, *Operations Research*, 65, 1, 97-103.

- [26] Hey, J.D., and C. Orme, (1994), Investigating generalizations of expected utility theory using experimental data, *Econometrica*, 1291-1326.
- [27] Hirsch, F., C. Profetta, B. Roynette and M. Yor, (2011), Constructing self-similar martingales via two Skorokhod embeddings, *Sem. de Prob. XLIII* 451-503 LNM 2006, Springer-Verlag, Berlin.
- [28] Ingersoll, J.E. and L. Jin, (2013), Realization Utility with Reference-Dependent Preferences, *Review of Financial Studies*, 26 (3), 723-767.
- [29] Johnson, E.J. and R. Ratcliff, (2013), Computational and Process Models of Decision Making in Psychology and behavioural Economics, in *Neuroeconomics: Decision Making and the Brain*. ed. by P.W. Glimcher and E. Fehr, Academic Press, 35-48.
- [30] Kahneman, D. and A. Tversky, (1979), Prospect Theory: An Analysis of Decision Under Risk, *Econometrica*, 46, 171-185.
- [31] Karatzas, I. and S. E. Shreve, (1991), *Brownian Motion and Stochastic Calculus. (Second Edition)*, Springer.
- [32] Karni, E. and Safra Z., (1990), behaviourally Consistent Optimal Stopping Rules, *Journal of Economic Theory*, 51(2), 391-402.
- [33] Kyle, A.S., Ou-Yang, H. and W. Xiong, (2006), Prospect Theory and Liquidation Decisions, *Journal of Economic Theory*, 129, 273-288.
- [34] Lempa, J, (2012), Optimal stopping with information constraint, *Appl. Math. Optim.*, 66(2) p147-173.
- [35] Liang, G. and W. Wei, (2016), Optimal Switching at Poisson Random Intervention Times, *Discrete and Continuous Dynamical Systems-Series B*, 21(5), 1483-1505
- [36] Luce, R.D., (1959), *Individual Choice behaviour: A Theoretical Analysis*, Wiley, New York.
- [37] Loomes, G. and R. Sugden, (1982), Regret theory: An alternative theory of rational choice under uncertainty, *Economic Journal*, 92, 805-824.
- [38] Maccheroni, F., (2002), Maxmin under Risk, *Economic Theory*, 19, 823-831.
- [39] Machina, M., (1985), Stochastic Choice Functions Generated from Deterministic Preferences over Lotteries, *Economic Journal*, 95, 379, 575-594.

- [40] Magnani, J., (2016), Reference-Dependent Preferences and Optimal Stopping Decisions: A Laboratory Investigation, Working paper.
- [41] McKean, H.P., (1965), Appendix: A free boundary problem for the heat equation arising from a problem in mathematical economics, *Industrial Manag. Review*, 6 p32–39.
- [42] Menaldi, J.L. and M. Robin, (2016), On Some Optimal Stopping Problems with Constraint, *SIAM Journal on Control and Optimization*, 54(5), 2650-2671
- [43] Oprea, R., D. Friedman, and S.T. Anderson, (2009), Learning to wait: A laboratory investigation, *The Review of Economic Studies*, 76(3), 1103-1124.
- [44] Pedersen, J. and G. Pekir, (2001), The Azema-Yor embedding in non-singular diffusions, *Stochastic Process. Appl.*, 96, 305-312.
- [45] Peskir, G. and A. Shiryaev, (2006), *Optimal Stopping and Free-Boundary Problems*, Birkhauser Verlag, Basel.
- [46] Petrovski, I.G., (1966), *Ordinary differential equations* (Revised English Edition), Prentice-Hall, New Jersey
- [47] Pham, H., (2009), *Continuous-time Stochastic Control and Optimization with Financial Applications*, Springer
- [48] Pham, H., and P. Tankov, (2008), A model of consumption under liquidity risk with random trading times, *Mathematical Finance*, 18(4) p613–627.
- [49] Quiggin, J., (1982), A Theory of Anticipated Utility, *Journal of Economic Behaviour and Organisation*, 3, 323-343.
- [50] Rogers, L.C.G., (2001), The Relaxed Investor and Parameter Uncertainty, *Finance and Stochastics*, 5, 131-154.
- [51] Rogers, L.C.G. and D. Williams, (2000), *Diffusions, Markov Processes and Martingales: Itô Calculus* Wiley, Chichester.
- [52] Rogers, L.C.G. and O. Zane, (2000), A simple model of liquidity effects, *Advances in Finance and Stochastics*, Eds: K. Sandmann and P. Schönbucher, Springer, Berlin. p161–176.
- [53] Rogozin, B.A., (1996), On the distribution of functionals related to boundary problems for processes with independent increments, *Th. Prob. Appl.*, 11, 580-591.

- [54] Sannikov, Y, (2008), A continuous-time version of the Principi-Pl-Agent problem, *Review of Economic Studies*, 75(3) p957–984.
- [55] Scheinkman, J. A. and W. Xiong, (2003), Overconfidence and Speculative Bubbles, *Journal of Political Economy*, 111, 6, 1183-1220.
- [56] Skorokhod, A.V., (1965), *Studies in the theory of random processes*, Addison-Wesley, Reading, Mass..
- [57] Strack, P. and P. Vieffers, (2017), Too Proud to Stop: Regret in Dynamic Decisions, *SSRN Working paper*, id2465840.
- [58] Tversky, A., (1969), Intransitivity of preferences, *Psychological review*, 76, 31.
- [59] Tversky, A. and D. Kahneman, (1992), Advances in Prospect Theory: Cumulative Representation of Uncertainty, *Journal of Risk and Uncertainty*, 5, 297-323.
- [60] Wakker P., (1994), Separating marginal utility and probabilistic risk aversion, *Theory and Decision*, 36(1), 1-44.
- [61] Wakker P., 2010, Prospect Theory for Risk and Ambiguity, *Cambridge University Press*.
- [62] Xu, Z.Q. and X.Y. Zhou, (2013), Optimal stopping under probability distortion. *Ann. Appl. Prob.*, 23, 1, 251-282.